Mon Dec 3 08:02:29 2001

frag1.rge

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: November 29, 2001, 12:28:27 ; Search time 1391.6 Seconds
(without alignments)
94.839 Million cell updates/sec

Title: FRAG1
Perfect score: 8
Sequence: 1 AACGTTCG 8

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 661134

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vi:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rod:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rod:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	Query Match	Length	DB	ID	Description
c	1	8	100.0	8	6	AX104477	AX104477 Sequence
	2	8	100.0	14	6	AR148617	AR148617 Sequence
	3	8	100.0	20	6	AR053410	AR053410 Sequence
c	4	8	100.0	20	6	AR146299	AR146299 Sequence
c	5	8	100.0	20	6	AX105150	AX105150 Sequence
	6	8	100.0	20	6	I84266	I84266 Sequence 37
	7	8	100.0	22	6	AX036945	AX036945 Sequence
	8	8	100.0	22	6	AX046993	AX046993 Sequence
	9	8	100.0	22	6	AX083675	AX083675 Sequence
	10	8	100.0	22	6	AX083676	AX083676 Sequence
	11	8	100.0	22	6	AX135650	AX135650 Sequence
	12	8	100.0	22	6	AX148636	AX148636 Sequence
	13	8	100.0	22	6	AX148637	AX148637 Sequence
	14	8	100.0	22	6	AX174913	AX174913 Sequence
c	15	8	100.0	23	6	AX083677	AX083677 Sequence
c	16	8	100.0	23	6	AX148638	AX148638 Sequence
	17	8	100.0	24	6	A84219	A84219 Sequence 14
	18	8	100.0	24	6	A84221	A84221 Sequence 16
c	19	8	100.0	24	6	AR106711	AR106711 Sequence
c	20	8	100.0	24	6	AX034754	AX034754 Sequence
	21	8	100.0	24	6	AX036501	AX036501 Sequence
c	22	8	100.0	24	6	I24359	I24359 Sequence 9
	23	8	100.0	26	6	AX083679	AX083679 Sequence
	24	8	100.0	26	6	AX148640	AX148640 Sequence
	25	8	100.0	27	6	AX035610	AX035610 Sequence
c	26	8	100.0	27	6	AX035610	AX035610 Sequence
	27	8	100.0	30	6	AR001144	AR001144 Sequence
	28	8	100.0	30	6	AR003022	AR003022 Sequence
	29	8	100.0	30	6	AR032996	AR032996 Sequence
c	30	8	100.0	30	6	AR122000	AR122000 Sequence
	31	8	100.0	30	6	I29736	I29736 Sequence 60
c	32	8	100.0	30	6	I71906	I71906 Sequence 10
	33	8	100.0	30	6	I76866	I76866 Sequence 8
	34	8	100.0	30	6	I87818	I87818 Sequence 8
	35	8	100.0	30	6	I91410	I91410 Sequence 60
	36	8	100.0	31	6	AR126483	AR126483 Sequence
	37	8	100.0	31	6	AR126484	AR126484 Sequence
	38	8	100.0	31	6	AR126489	AR126489 Sequence
	39	8	100.0	31	6	AR126495	AR126495 Sequence
c	40	8	100.0	36	6	AX093421	AX093421 Sequence
	41	8	100.0	39	6	AR011688	AR011688 Sequence
	42	8	100.0	39	6	I58330	I58330 Sequence 1
	43	8	100.0	39	6	I92478	I92478 Sequence 1
c	44	8	100.0	39	8	SJU82656	U82656 Scybalium j
c	45	8	100.0	40	6	AR035187	AR035187 Sequence

ALIGNMENTS

RESULT 1
AX104477/c 8 bp DNA PAT 30-APR-2001
LOCUS AX104477
DEFINITION Sequence 669 from Patent WO0122972.
ACCESSION AX104477
VERSION AX104477.1 GI:13920674
KEYWORDS .
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 8)
AUTHORS Krieg,A.M., Schetter,C. and Vollmer,J.C.
TITLE Immunostimulatory nucleic acids
JOURNAL Patent: WO 0122972-A 669 05-APR-2001;
UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical GmbH (DE)

FEATURES Location/Qualifiers
source 1. .8
/organism="synthetic construct"
/db_xref="taxon:32630"

```
BASE COUNT         2 a      2 c      2 g      2 t
ORIGIN

  Query Match                 100.0%;  Score 8;  DB 6;  Length 8;
  Best Local Similarity  100.0%;  Pred. No. 2.1e+09;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 AACGTTCG 8
          ||||||||
Db      8 AACGTTCG 1


RESULT   2
AR148617
LOCUS       AR148617        14 bp    DNA                PAT       08-AUG-2001
DEFINITION  Sequence 11 from patent US 6225292.
ACCESSION   AR148617
VERSION     AR148617.1  GI:15112707
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 14)
  AUTHORS   Raz,E. and Roman,M.
  TITLE     Inhibitors of DNA immunostimulatory sequence activity
  JOURNAL   Patent: US 6225292-A 11 01-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .14
                     /organism="unknown"
BASE COUNT         4 a      4 c      2 g      4 t
ORIGIN

  Query Match                 100.0%;  Score 8;  DB 6;  Length 14;
  Best Local Similarity  100.0%;  Pred. No. 1.1e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 AACGTTCG 8
          ||||||||
Db      6 AACGTTCG 13


RESULT   3
AR053410/c
LOCUS       AR053410        20 bp    DNA                PAT       29-SEP-1999
DEFINITION  Sequence 11 from patent US 5834245.
ACCESSION   AR053410
VERSION     AR053410.1  GI:5978272
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Nakamura,Y. and Fujiwara,Y.
  TITLE     PRLTS proteins and DNA's encoding the same
  JOURNAL   Patent: US 5834245-A 11 10-NOV-1998;
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="unknown"
BASE COUNT         6 a      5 c      7 g      2 t
ORIGIN

  Query Match                 100.0%;  Score 8;  DB 6;  Length 20;
  Best Local Similarity  100.0%;  Pred. No. 1e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 AACGTTCG 8
          ||||||||
Db     17 AACGTTCG 10


RESULT   4
AR146299/c
LOCUS       AR146299        20 bp    DNA                PAT       08-AUG-2001
DEFINITION  Sequence 11 from patent US 6218371.
ACCESSION   AR146299
VERSION     AR146299.1  GI:15109488
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Krieg,A.M. and Weiner,G.
  TITLE     Methods and products for stimulating the immune system using
            immunotherapeutic oligonucleotides and cytokines
  JOURNAL   Patent: US 6218371-A 11 17-APR-2001;
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="unknown"
BASE COUNT         4 a      7 c      3 g      6 t
ORIGIN

  Query Match                 100.0%;  Score 8;  DB 6;  Length 20;
  Best Local Similarity  100.0%;  Pred. No. 1e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 AACGTTCG 8
          ||||||||
Db     17 AACGTTCG 10


RESULT   5
AX105150/c
LOCUS       AX105150        20 bp    DNA                PAT       30-APR-2001
DEFINITION  Sequence 48 from Patent WO0122990.
ACCESSION   AX105150
VERSION     AX105150.1  GI:13921300
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Hartmann,G.D., Bratzler,R.L. and Krieg,A.U.
  TITLE     Methods related to immunostimulatory nucleic acid-induced
            interferon
  JOURNAL   Patent: WO 0122990-A 48 05-APR-2001;
            Coley Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH
            FOUNDATION (US)
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Synthetic Oligonucleotide"
BASE COUNT         4 a      7 c      3 g      6 t
ORIGIN

  Query Match                 100.0%;  Score 8;  DB 6;  Length 20;
  Best Local Similarity  100.0%;  Pred. No. 1e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 AACGTTCG 8
          ||||||||
Db     17 AACGTTCG 10


RESULT   6
I84266
LOCUS       I84266          20 bp    DNA                PAT       04-APR-1998
DEFINITION  Sequence 37 from patent US 5695926.
ACCESSION   I84266
VERSION     I84266.1  GI:3021786
KEYWORDS    .
```

```
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Cros,P., Allibert,P., Mallet,F., Mabilat,C. and Mandrand,B.
  TITLE     Sandwich hybridization assays using very short capture probes
            noncovalently bound to a hydrophobic support
  JOURNAL   Patent: US 5695926-A 37 09-DEC-1997;
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="unknown"
BASE COUNT         4 a      5 c      6 g      5 t
ORIGIN

  Query Match             100.0%;  Score 8;  DB 6;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       5 AACGTTCG 12


RESULT   7
AX036945                                                PAT       16-NOV-2000
LOCUS       AX036945        22 bp    DNA
DEFINITION  Sequence 2 from Patent FR2790955.
ACCESSION   AX036945
VERSION     AX036945.1  GI:11226373
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Carpentier,A.
  JOURNAL   Patent: FR 2790955-A 2 22-SEP-2000;
            ASSIST PUBL HOPITAUX DE PARIS (FR)
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note=" oligodesoxynucleotide"
BASE COUNT         6 a      3 c      7 g      6 t
ORIGIN

  Query Match             100.0%;  Score 8;  DB 6;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       9 AACGTTCG 16


RESULT   8
AX046993                                                PAT       15-DEC-2000
LOCUS       AX046993        22 bp    DNA
DEFINITION  Sequence 2 from Patent WO0067787.
ACCESSION   AX046993
VERSION     AX046993.1  GI:11876420
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Moss,R.B.
  TITLE     Hiv immunogenic compositions and methods
  JOURNAL   Patent: WO 0067787-A 2 16-NOV-2000;
            THE IMMUNE RESPONSE CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1  22
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="phosphorothioate-modified synthetic
                     oligodeoxynucleotide"
BASE COUNT         6 a      3 c      7 g      6 t
ORIGIN

  Query Match             100.0%;  Score 8;  DB 6;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       9 AACGTTCG 16


RESULT   9
AX083675                                                PAT       28-FEB-2001
LOCUS       AX083675        22 bp    DNA
DEFINITION  Sequence 1 from Patent WO0112223.
ACCESSION   AX083675
VERSION     AX083675.1  GI:13185407
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   van Nest,G.
  TITLE     Methods of modulating an immune response using immunostimulatory s
            equences and compositions for use therein
  JOURNAL   Patent: WO 0112223-A 1 22-FEB-2001;
            Dynavax Technologies Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Synthetic construct"
BASE COUNT         6 a      3 c      7 g      6 t
ORIGIN

  Query Match             100.0%;  Score 8;  DB 6;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       9 AACGTTCG 16


RESULT   10
AX083676                                                PAT       28-FEB-2001
LOCUS       AX083676        22 bp    DNA
DEFINITION  Sequence 2 from Patent WO0112223.
ACCESSION   AX083676
VERSION     AX083676.1  GI:13185408
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   van Nest,G.
  TITLE     Methods of modulating an immune response using immunostimulatory s
            equences and compositions for use therein
  JOURNAL   Patent: WO 0112223-A 2 22-FEB-2001;
            Dynavax Technologies Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Synthetic construct"
BASE COUNT         6 a      4 c      7 g      5 t
```

ORIGIN

Query Match 100.0%; Score 8; DB 6; Length 22;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACGTTCG 8
||||||||
Db 9 AACGTTCG 16

RESULT 11
AX135650
LOCUS AX135650 22 bp DNA PAT 29-MAY-2001
DEFINITION Sequence 21 from Patent WO0132877.
ACCESSION AX135650
VERSION AX135650.1 GI:14271920
KEYWORDS .
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 22)
AUTHORS Mackichan,M.L.
TITLE Cpg receptor (cpg-r) and methods relating thereto
JOURNAL Patent: WO 0132877-A 21 10-MAY-2001;
CHIRON CORPORATION (US)
FEATURES Location/Qualifiers
source 1. .22
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="CpG oligonucleotide"
BASE COUNT 6 a 3 c 7 g 6 t
ORIGIN

Query Match 100.0%; Score 8; DB 6; Length 22;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACGTTCG 8
||||||||
Db 9 AACGTTCG 16

RESULT 12
AX148636
LOCUS AX148636 22 bp DNA PAT 08-JUN-2001
DEFINITION Sequence 1 from Patent WO0135991.
ACCESSION AX148636
VERSION AX148636.1 GI:14347254
KEYWORDS .
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 22)
AUTHORS Tuck,S. and van Nest,G.
TITLE Immunomodulatory compositions containing an immunostimulatory
JOURNAL sequence linked to antigen and methods of use thereof
Patent: WO 0135991-A 1 25-MAY-2001;
FEATURES Dynavax Technologies Corporation (US)
source Location/Qualifiers
1. .22
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="synthetic construct"
BASE COUNT 6 a 3 c 7 g 6 t
ORIGIN

Query Match 100.0%; Score 8; DB 6; Length 22;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACGTTCG 8
||||||||
Db 9 AACGTTCG 16

RESULT 13
AX148637
LOCUS AX148637 22 bp DNA PAT 08-JUN-2001
DEFINITION Sequence 2 from Patent WO0135991.
ACCESSION AX148637
VERSION AX148637.1 GI:14347255
KEYWORDS .
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 22)
AUTHORS Tuck,S. and van Nest,G.
TITLE Immunomodulatory compositions containing an immunostimulatory
JOURNAL sequence linked to antigen and methods of use thereof
Patent: WO 0135991-A 2 25-MAY-2001;
FEATURES Dynavax Technologies Corporation (US)
source Location/Qualifiers
1. .22
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="synthetic construct"
BASE COUNT 6 a 4 c 7 g 5 t
ORIGIN

Query Match 100.0%; Score 8; DB 6; Length 22;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACGTTCG 8
||||||||
Db 9 AACGTTCG 16

RESULT 14
AX174913
LOCUS AX174913 22 bp DNA PAT 03-JUL-2001
DEFINITION Sequence 1 from Patent WO0143778.
ACCESSION AX174913
VERSION AX174913.1 GI:14598409
KEYWORDS .
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 22)
AUTHORS Felgner,P.L. and Zelphati,O.
TITLE Use of cationic lipids for intracellular protein delivery
JOURNAL Patent: WO 0143778-A 1 21-JUN-2001;
FEATURES Gene Therapy Systems, Inc. (US)
source Location/Qualifiers
1. .22
/organism="synthetic construct"
/db_xref="taxon:32630"
modified_base /note="Synthetic peptide"
1
/note="n=T-NH2"
modified_base /mod_base=OTHER
22
/note="n=A-Rhodamine"
BASE COUNT /mod_base=OTHER
5 a 3 c 7 g 5 t 2 others
ORIGIN

Query Match 100.0%; Score 8; DB 6; Length 22;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0, Gaps 0,

```
Qy       1 AACGTTCG 8
           ||||||||
Db       9 AACGTTCG 16

RESULT  15
AX083677/c                                          PAT       28-FEB-2001
LOCUS       AX083677        23 bp    DNA
DEFINITION  Sequence 3 from Patent WO0112223.
ACCESSION   AX083677
VERSION     AX083677.1  GI:13185409
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 23)
  AUTHORS   van Nest,G.
  TITLE     Methods of modulating an immune response using immunostimulatory s
            equences and compositions for use therein
  JOURNAL   Patent: WO 0112223-A 3 22-FEB-2001;
            Dynavax Technologies Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .23
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Synthetic construct"
BASE COUNT         6 a      8 c      3 g      6 t
ORIGIN

  Query Match           100.0%;  Score 8;  DB 6;  Length 23;
  Best Local Similarity 100.0%;  Pred. No. 1e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      14 AACGTTCG 7

RESULT  16
AX148638/c                                          PAT       08-JUN-2001
LOCUS       AX148638        23 bp    DNA
DEFINITION  Sequence 3 from Patent WO0135991.
ACCESSION   AX148638
VERSION     AX148638.1  GI:14347256
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 23)
  AUTHORS   Tuck,S. and van Nest,G.
  TITLE     Immunomodulatory compositions containing an immunostimulatory
            sequence linked to antigen and methods of use thereof
  JOURNAL   Patent: WO 0135991-A 3 25-MAY-2001;
            Dynavax Technologies Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .23
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic construct"
BASE COUNT         6 a      8 c      3 g      6 t
ORIGIN

  Query Match           100.0%;  Score 8;  DB 6;  Length 23;
  Best Local Similarity 100.0%;  Pred. No. 1e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      14 AACGTTCG 7

RESULT  17
A84219                                              PAT       21-JAN-2000
LOCUS       A84219          24 bp    DNA
DEFINITION  Sequence 14 from Patent WO9846733.
ACCESSION   A84219
VERSION     A84219.1  GI:6733267
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   Tybulewicz,V. and Martin,J.
  TITLE     CHROMOSOME TRANSFER (XMMCT) TO ES CELLS, INDUCED DURING THE
            EXPOSURE OF MICROCELLS TO RADIATION
  JOURNAL   Patent: WO 9846733-A 14 22-OCT-1998;
            TYBULEWICZ VICTOR (GB); MARTIN JOANNE (GB)
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT         6 a      5 c      5 g      8 t
ORIGIN

  Query Match           100.0%;  Score 8;  DB 6;  Length 24;
  Best Local Similarity 100.0%;  Pred. No. 1e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       1 AACGTTCG 8

RESULT  18
A84221                                              PAT       21-JAN-2000
LOCUS       A84221          24 bp    DNA
DEFINITION  Sequence 16 from Patent WO9846733.
ACCESSION   A84221
VERSION     A84221.1  GI:6733269
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   Tybulewicz,V. and Martin,J.
  TITLE     CHROMOSOME TRANSFER (XMMCT) TO ES CELLS, INDUCED DURING THE
            EXPOSURE OF MICROCELLS TO RADIATION
  JOURNAL   Patent: WO 9846733-A 16 22-OCT-1998;
            TYBULEWICZ VICTOR (GB); MARTIN JOANNE (GB)
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT         6 a      5 c      5 g      8 t
ORIGIN

  Query Match           100.0%;  Score 8;  DB 6;  Length 24;
  Best Local Similarity 100.0%;  Pred. No. 1e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       1 AACGTTCG 8

RESULT  19
AR106711/c                                          PAT       14-FEB-2001
LOCUS       AR106711        24 bp    DNA
DEFINITION  Sequence 9 from patent US 6107087.
ACCESSION   AR106711
VERSION     AR106711.1  GI:12821241
```

```
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   O'Neill,G.P. and Mancini,J.A.
  TITLE     High level expression of human cyclooxygenase-2
  JOURNAL   Patent: US 6107087-A 9 22-AUG-2000;
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="unknown"
BASE COUNT         5 a       7 c       4 g       8 t
ORIGIN
```

```
Query Match                 100.0%;  Score 8;  DB 6;  Length 24;
Best Local Similarity  100.0%;  Pred. No. 1e+05;
Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 AACGTTCG 8
          ||||||||
Db     23 AACGTTCG 16
```

```
RESULT  20
AX034754/c
LOCUS       AX034754         24 bp    DNA                     PAT       15-NOV-2000
DEFINITION  Sequence 36 from Patent WO0052203.
ACCESSION   AX034754
VERSION     AX034754.1  GI:11190713
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   Anthony,R.M., Brown,T.J. and French,G.L.
  TITLE     Identification of bacteria
  JOURNAL   Patent: WO 0052203-A 36 08-SEP-2000;
            ANTHONY RICHARD MICHAEL (GB) ; BROWN TIMOTHY JAMES (GB) ; KING S
            COLLEGE LONDON (GB) ; FRENCH GARY LAWRENCE (GB) ; GUY S & ST THOMAS
            S NATIONAL H (GB)
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="OLIGO"
BASE COUNT         4 a       5 c       8 g       7 t
ORIGIN
```

```
Query Match                 100.0%;  Score 8;  DB 6;  Length 24;
Best Local Similarity  100.0%;  Pred. No. 1e+05;
Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 AACGTTCG 8
          ||||||||
Db     22 AACGTTCG 15
```

```
RESULT  21
AX036501
LOCUS       AX036501         24 bp    DNA                     PAT       16-NOV-2000
DEFINITION  Sequence 163 from Patent DE19915141.
ACCESSION   AX036501
VERSION     AX036501.1  GI:11226111
KEYWORDS    .
SOURCE      Pseudomonas fluorescens.
  ORGANISM  Pseudomonas fluorescens
            Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
            Pseudomonas.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   Krupp,G.
  JOURNAL   Patent: DE 19915141-A 163 28-SEP-2000;
            ARTUS GES FUER MOLEKULARBIOLOG (DE)
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="Pseudomonas fluorescens"
                     /db_xref="taxon:294"
BASE COUNT         9 a       5 c       6 g       4 t
ORIGIN
```

```
Query Match                 100.0%;  Score 8;  DB 6;  Length 24;
Best Local Similarity  100.0%;  Pred. No. 1e+05;
Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 AACGTTCG 8
          ||||||||
Db      3 AACGTTCG 10
```

```
RESULT  22
I24359/c
LOCUS       I24359           24 bp    DNA                     PAT       07-OCT-1996
DEFINITION  Sequence 9 from patent US 5543297.
ACCESSION   I24359
VERSION     I24359.1  GI:1604229
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   Cromlish,W.A., Kennedy,B.P., O'Neill,G., Vickers,P.J., Wong,E. and
            Mancini,J.A.
  TITLE     Human cyclooxygenase-2 cDNA and assays for evaluating
            cyclooxygenase-2 activity
  JOURNAL   Patent: US 5543297-A 9 06-AUG-1996;
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="unknown"
BASE COUNT         5 a       7 c       4 g       8 t
ORIGIN
```

```
Query Match                 100.0%;  Score 8;  DB 6;  Length 24;
Best Local Similarity  100.0%;  Pred. No. 1e+05;
Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 AACGTTCG 8
          ||||||||
Db     23 AACGTTCG 16
```

```
RESULT  23
AX083679
LOCUS       AX083679         26 bp    DNA                     PAT       28-FEB-2001
DEFINITION  Sequence 5 from Patent WO0112223.
ACCESSION   AX083679
VERSION     AX083679.1  GI:13185411
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 26)
  AUTHORS   van Nest,G.
  TITLE     Methods of modulating an immune response using immunostimulatory s
            equences and compositions for use therein
  JOURNAL   Patent: WO 0112223-A 5 22-FEB-2001;
            Dynavax Technologies Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .26
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Synthetic construct"
BASE COUNT         5 a       9 c       4 g       8 t
ORIGIN
```

```
Query Match             100.0%;  Score 8;  DB 6;  Length 26;
Best Local Similarity   100.0%;  Pred. No. 1e+05;
Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       6 AACGTTCG 13


RESULT  24
AX148640
LOCUS       AX148640        26 bp    DNA                     PAT       08-JUN-2001
DEFINITION  Sequence 5 from Patent WO0135991.
ACCESSION   AX148640
VERSION     AX148640.1  GI:14347258
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 26)
  AUTHORS   Tuck,S. and van Nest,G.
  TITLE     Immunomodulatory compositions containing an immunostimulatory
            sequence linked to antigen and methods of use thereof
  JOURNAL   Patent: WO 0135991-A 5 25-MAY-2001;
            Dynavax Technologies Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .26
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic construct"
BASE COUNT         5 a      9 c      4 g      8 t
ORIGIN

Query Match             100.0%;  Score 8;  DB 6;  Length 26;
Best Local Similarity   100.0%;  Pred. No. 1e+05;
Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       6 AACGTTCG 13


RESULT  25
AX035610
LOCUS       AX035610        27 bp    DNA                     PAT       15-NOV-2000
DEFINITION  Sequence 25 from Patent WO0052152.
ACCESSION   AX035610
VERSION     AX035610.1  GI:11191205
KEYWORDS    .
SOURCE      Bacillus subtilis.
  ORGANISM  Bacillus subtilis
            Bacteria; Firmicutes; Bacillus/Clostridium group;
            Bacillus/Staphylococcus group; Bacillus.
REFERENCE   1  (bases 1 to 27)
  AUTHORS   Stachelhaus,T., Konz,D., Mootz,H. and Marahiel,M.A.
  TITLE     Non-ribosomal peptide synthetases, method for producing same and
            the use thereof
  JOURNAL   Patent: WO 0052152-A 25 08-SEP-2000;
            STACHELHAUS TORSTEN (DE) ; KONZ DIRK (DE) ; MOOTZ HENNING (DE) ;
            MARAHIEL MOHAMED A (DE)
FEATURES             Location/Qualifiers
     source          1. .27
                     /organism="Bacillus subtilis"
                     /db_xref="taxon:1423"
BASE COUNT         5 a      7 c      8 g      7 t
ORIGIN

Query Match             100.0%;  Score 8;  DB 6;  Length 27;
Best Local Similarity   100 0%;  Pred. No. 1e+05;
Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      12 AACGTTCG 19


RESULT  26
AX035610/c
LOCUS       AX035610        27 bp    DNA                     PAT       15-NOV-2000
DEFINITION  Sequence 25 from Patent WO0052152.
ACCESSION   AX035610
VERSION     AX035610.1  GI:11191205
KEYWORDS    .
SOURCE      Bacillus subtilis.
  ORGANISM  Bacillus subtilis
            Bacteria; Firmicutes; Bacillus/Clostridium group;
            Bacillus/Staphylococcus group; Bacillus.
REFERENCE   1  (bases 1 to 27)
  AUTHORS   Stachelhaus,T., Konz,D., Mootz,H. and Marahiel,M.A.
  TITLE     Non-ribosomal peptide synthetases, method for producing same and
            the use thereof
  JOURNAL   Patent: WO 0052152-A 25 08-SEP-2000;
            STACHELHAUS TORSTEN (DE) ; KONZ DIRK (DE) ; MOOTZ HENNING (DE) ;
            MARAHIEL MOHAMED A (DE)
FEATURES             Location/Qualifiers
     source          1. .27
                     /organism="Bacillus subtilis"
                     /db_xref="taxon:1423"
BASE COUNT         5 a      7 c      8 g      7 t
ORIGIN

Query Match             100.0%;  Score 8;  DB 6;  Length 27;
Best Local Similarity   100.0%;  Pred. No. 1e+05;
Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      17 AACGTTCG 10


RESULT  27
AR001144
LOCUS       AR001144        30 bp    DNA                     PAT       04-DEC-1998
DEFINITION  Sequence 8 from patent US 5738990.
ACCESSION   AR001144
VERSION     AR001144.1  GI:3963211
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Edwards,C.A., Fry,K.E., Cantor,C.R. and Andrews,B.M.
  TITLE     Sequence-directed DNA-binding molecules compositions and methods
  JOURNAL   Patent: US 5738990-A 8 14-APR-1998;
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="unknown"
BASE COUNT         16 a      4 c      6 g      4 t
ORIGIN

Query Match             100.0%;  Score 8;  DB 6;  Length 30;
Best Local Similarity   100.0%;  Pred. No. 1e+05;
Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       8 AACGTTCG 15


RESULT  28
```

```
AR003022
LOCUS       AR003022          30 bp    DNA                PAT       04-DEC-1998
DEFINITION  Sequence 8 from patent US 5744131.
ACCESSION   AR003022
VERSION     AR003022.1  GI:3964281
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Edwards,C.A., Fry,K.E., Cantor,C.R. and Andrews,B.M.
  TITLE     Sequence-directed DNA-binding molecules compositions and methods
  JOURNAL   Patent: US 5744131-A 8 28-APR-1998;
FEATURES              Location/Qualifiers
     source           1. .30
                      /organism="unknown"
BASE COUNT         16 a       4 c       6 g       4 t
ORIGIN
```

```
Query Match                   100.0%;  Score 8;  DB 6;  Length 30;
Best Local Similarity  100.0%;  Pred. No. 1e+05;
Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       8 AACGTTCG 15
```

```
RESULT  29
AR032996
LOCUS       AR032996          30 bp    DNA                PAT       29-SEP-1999
DEFINITION  Sequence 608 from patent US 5869241.
ACCESSION   AR032996
VERSION     AR032996.1  GI:5948601
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
  TITLE     Method of determining DNA sequence preference of a DNA-binding
            molecule
  JOURNAL   Patent: US 5869241-A 608 09-FEB-1999;
FEATURES              Location/Qualifiers
     source           1. .30
                      /organism="unknown"
BASE COUNT         16 a       4 c       6 g       4 t
ORIGIN
```

```
Query Match                   100.0%;  Score 8;  DB 6;  Length 30;
Best Local Similarity  100.0%;  Pred. No. 1e+05;
Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       8 AACGTTCG 15
```

```
RESULT  30
AR122000/c
LOCUS       AR122000          30 bp    DNA                PAT       16-MAY-2001
DEFINITION  Sequence 19 from patent US 6160203.
ACCESSION   AR122000
VERSION     AR122000.1  GI:14105576
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Ferri,S. and Toguri,T.
  TITLE     DNA strands coding for glycerol-e-phosphate acyltransferase
  JOURNAL   Patent: US 6160203-A 19 12-DEC-2000;
FEATURES              Location/Qualifiers
     source           1. .30
                      /organism="unknown"
BASE COUNT         9 a       8 c       6 g       7 t
ORIGIN
```

```
Query Match                   100.0%;  Score 8;  DB 6;  Length 30;
Best Local Similarity  100.0%;  Pred. No. 1e+05;
Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      24 AACGTTCG 17
```

```
RESULT  31
I29736
LOCUS       I29736            30 bp    DNA                PAT       06-FEB-1997
DEFINITION  Sequence 608 from patent US 5578444.
ACCESSION   I29736
VERSION     I29736.1  GI:1820527
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
  TITLE     Sequence-directed DNA-binding molecules compositions and methods
  JOURNAL   Patent: US 5578444-A 608 26-NOV-1996;
FEATURES              Location/Qualifiers
     source           1. .30
                      /organism="unknown"
BASE COUNT         16 a       4 c       6 g       4 t
ORIGIN
```

```
Query Match                   100.0%;  Score 8;  DB 6;  Length 30;
Best Local Similarity  100.0%;  Pred. No. 1e+05;
Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       8 AACGTTCG 15
```

```
RESULT  32
I71906/c
LOCUS       I71906            30 bp    DNA                PAT       03-APR-1998
DEFINITION  Sequence 10 from patent US 5683868.
ACCESSION   I71906
VERSION     I71906.1  GI:3008045
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   LaRossa,R.Alan, Majarian,W.Robert and Van Dyk,T.Kangas.
  TITLE     Highly sensitive method for detecting environmental insults
  JOURNAL   Patent: US 5683868-A 10 04-NOV-1997;
FEATURES              Location/Qualifiers
     source           1. .30
                      /organism="unknown"
BASE COUNT         9 a       8 c       8 g       5 t
ORIGIN
```

```
Query Match                   100.0%;  Score 8;  DB 6;  Length 30;
Best Local Similarity  100.0%;  Pred. No. 1e+05;
Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
```

Mon Dec 3 08:02:29 2001

```
          ||||||||
Db      19 AACGTTCG 12


RESULT  33
I76866
LOCUS       I76866          30 bp    DNA                PAT       03-APR-1998
DEFINITION  Sequence 8 from patent US 5693463.
ACCESSION   I76866
VERSION     I76866.1  GI:3013020
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Edwards,C.A., Fry,K.E., Cantor,C.R. and Andrews,B.M.
  TITLE     Method of ordering sequence binding preferences of a DNA-binding
            molecule
  JOURNAL   Patent: US 5693463-A 8 02-DEC-1997;
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="unknown"
BASE COUNT        16 a      4 c      6 g      4 t
ORIGIN

  Query Match           100.0%;  Score 8;  DB 6;  Length 30;
  Best Local Similarity 100.0%;  Pred. No. 1e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 AACGTTCG 8
          ||||||||
Db      8 AACGTTCG 15


RESULT  34
I87818
LOCUS       I87818          30 bp    DNA                PAT       10-AUG-1998
DEFINITION  Sequence 8 from patent US 5716780.
ACCESSION   I87818
VERSION     I87818.1  GI:3407758
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Edwards,C.A., Fry,K.E., Cantor,C.R. and Andrews,B.M.
  TITLE     Method of constructing sequence-specific DNA-binding molecules
  JOURNAL   Patent: US 5716780-A 8 10-FEB-1998;
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="unknown"
BASE COUNT        16 a      4 c      6 g      4 t
ORIGIN

  Query Match           100.0%;  Score 8;  DB 6;  Length 30;
  Best Local Similarity 100.0%;  Pred. No. 1e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 AACGTTCG 8
          ||||||||
Db      8 AACGTTCG 15


RESULT  35
I91410
LOCUS       I91410          30 bp    DNA                PAT       01-DEC-1998
DEFINITION  Sequence 608 from patent US 5726014.
ACCESSION   I91410
VERSION     I91410.1  GI:3935880
KEYWORDS    .
SOURCE      Unknown
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Edwards,C.A., Cantor,C.R., Andrews,B.M. and Turin,L.M.
  TITLE     Screening assay for the detection of DNA-binding molecules
  JOURNAL   Patent: US 5726014-A 608 10-MAR-1998;
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="unknown"
BASE COUNT        16 a      4 c      6 g      4 t
ORIGIN

  Query Match           100.0%;  Score 8;  DB 6;  Length 30;
  Best Local Similarity 100.0%;  Pred. No. 1e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 AACGTTCG 8
          ||||||||
Db      8 AACGTTCG 15


RESULT  36
AR126483
LOCUS       AR126483        31 bp    DNA                PAT       16-MAY-2001
DEFINITION  Sequence 110 from patent US 6180341.
ACCESSION   AR126483
VERSION     AR126483.1  GI:14113076
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 31)
  AUTHORS   Iverson,B.L., Georgiou,G. and Burks,E.A.
  TITLE     In vitro scanning saturation mutagenesis of proteins
  JOURNAL   Patent: US 6180341-A 110 30-JAN-2001;
FEATURES             Location/Qualifiers
     source          1. .31
                     /organism="unknown"
BASE COUNT        10 a      9 c      6 g      6 t
ORIGIN

  Query Match           100.0%;  Score 8;  DB 6;  Length 31;
  Best Local Similarity 100.0%;  Pred. No. 9.9e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 AACGTTCG 8
          ||||||||
Db     24 AACGTTCG 31


RESULT  37
AR126484
LOCUS       AR126484        31 bp    DNA                PAT       16-MAY-2001
DEFINITION  Sequence 111 from patent US 6180341.
ACCESSION   AR126484
VERSION     AR126484.1  GI:14113077
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 31)
  AUTHORS   Iverson,B.L., Georgiou,G. and Burks,E.A.
  TITLE     In vitro scanning saturation mutagenesis of proteins
  JOURNAL   Patent: US 6180341-A 111 30-JAN-2001;
FEATURES             Location/Qualifiers
     source          1. .31
                     /organism="unknown"
BASE COUNT         9 a     10 c      6 g      6 t
ORIGIN
```

```
  Query Match                    100.0%;  Score 8;  DB 6;  Length 31;
  Best Local Similarity  100.0%;  Pred. No. 9.9e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
          ||||||||
Db      24 AACGTTCG 31


RESULT  38
AR126489
LOCUS       AR126489         31 bp    DNA                     PAT       16-MAY-2001
DEFINITION  Sequence 116 from patent US 6180341.
ACCESSION   AR126489
VERSION     AR126489.1  GI:14113082
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 31)
  AUTHORS   Iverson,B.L., Georgiou,G. and Burks,E.A.
  TITLE     In vitro scanning saturation mutagenesis of proteins
  JOURNAL   Patent: US 6180341-A 116 30-JAN-2001;
FEATURES             Location/Qualifiers
     source          1. .31
                     /organism="unknown"
BASE COUNT        10 a       9 c       6 g       6 t
ORIGIN


  Query Match                    100.0%;  Score 8;  DB 6;  Length 31;
  Best Local Similarity  100.0%;  Pred. No. 9.9e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
          ||||||||
Db      24 AACGTTCG 31


RESULT  39
AR126495
LOCUS       AR126495         31 bp    DNA                     PAT       16-MAY-2001
DEFINITION  Sequence 122 from patent US 6180341.
ACCESSION   AR126495
VERSION     AR126495.1  GI:14113088
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 31)
  AUTHORS   Iverson,B.L., Georgiou,G. and Burks,E.A.
  TITLE     In vitro scanning saturation mutagenesis of proteins
  JOURNAL   Patent: US 6180341-A 122 30-JAN-2001;
FEATURES             Location/Qualifiers
     source          1. .31
                     /organism="unknown"
BASE COUNT        11 a       9 c       5 g       6 t
ORIGIN


  Query Match                    100.0%;  Score 8;  DB 6;  Length 31;
  Best Local Similarity  100.0%;  Pred. No. 9.9e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
          ||||||||
Db      24 AACGTTCG 31


RESULT  40
AX093421/c
LOCUS
DEFINITION  AX093421         36 bp    DNA                     PAT       30-MAR-2001
            Sequence 3 from Patent WO0118195.
ACCESSION   AX093421
VERSION     AX093421.1  GI:13509871
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 36)
  AUTHORS   Francis,K.P., Contag,P.R. and Joh,D.J.
  TITLE     Luciferase expression cassettes and methods of use
  JOURNAL   Patent: WO 0118195-A 3 15-MAR-2001;
            Xenogen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .36
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Primer XAR"
BASE COUNT        10 a       7 c       10 g      9 t
ORIGIN


  Query Match                    100.0%;  Score 8;  DB 6;  Length 36;
  Best Local Similarity  100.0%;  Pred. No. 9.8e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
          ||||||||
Db      35 AACGTTCG 28


RESULT  41
AR011688
LOCUS       AR011688         39 bp    DNA                     PAT       04-DEC-1998
DEFINITION  Sequence 1 from patent US 5763167.
ACCESSION   AR011688
VERSION     AR011688.1  GI:3969678
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 39)
  AUTHORS   Conrad,M.J.
  TITLE     Applications of fluorescent N-nucleosides and fluorescent
            structural analogs of N-nucleosides
  JOURNAL   Patent: US 5763167-A 1 09-JUN-1998;
FEATURES             Location/Qualifiers
     source          1. .39
                     /organism="unknown"
BASE COUNT        9 a       11 c      11 g      8 t
ORIGIN


  Query Match                    100.0%;  Score 8;  DB 6;  Length 39;
  Best Local Similarity  100.0%;  Pred. No. 9.8e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
          ||||||||
Db       1 AACGTTCG 8


RESULT  42
I58330
LOCUS       I58330           39 bp    DNA                     PAT       07-OCT-1997
DEFINITION  Sequence 1 from patent US 5652099.
ACCESSION   I58330
VERSION     I58330.1  GI:2477568
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 39)
  AUTHORS   Conrad,M.J.
  TITLE     Probes comprising fluorescent nucleosides and uses thereof
```

```
  JOURNAL   Patent: US 5652099-A 1 29-JUL-1997;
FEATURES             Location/Qualifiers
     source          1. .39
                     /organism="unknown"
BASE COUNT         9 a    11 c    11 g     8 t
ORIGIN

  Query Match           100.0%;  Score 8;  DB 6;  Length 39;
  Best Local Similarity 100.0%;  Pred. No. 9.8e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       1 AACGTTCG 8

RESULT  43
I92478                                 PAT       01-DEC-1998
LOCUS       I92478         39 bp    DNA
DEFINITION  Sequence 1 from patent US 5728525.
ACCESSION   I92478
VERSION     I92478.1  GI:3936948
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 39)
  AUTHORS   Conrad,M.J.
  TITLE     Fluorescent universal nucleic acid end label
  JOURNAL   Patent: US 5728525-A 1 17-MAR-1998;
FEATURES             Location/Qualifiers
     source          1. .39
                     /organism="unknown"
BASE COUNT         9 a    11 c    11 g     8 t
ORIGIN

  Query Match           100.0%;  Score 8;  DB 6;  Length 39;
  Best Local Similarity 100.0%;  Pred. No. 9.8e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       1 AACGTTCG 8

RESULT  44
SJU82656/c                             PLN       16-JAN-1998
LOCUS       SJU82656       39 bp    DNA
DEFINITION  Scybalium jamaicense small subunit ribosomal RNA gene,
            mitochondrial gene for mitochondrial RNA, complete sequence,
            partial sequence representing helix 6.
ACCESSION   U82656
VERSION     U82656.1  GI:2697046
KEYWORDS    .
SOURCE      Scybalium jamaicense.
  ORGANISM  Mitochondrion Scybalium jamaicense
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Balanophoraceae; Scybalium.
REFERENCE   1  (bases 1 to 39)
  AUTHORS   Duff,R.J. and Nickrent,D.L.
  TITLE     Characterization of mitochondrial small-subunit ribosomal RNAs from
            holoparasitic plants
  JOURNAL   J. Mol. Evol. 45 (6), 631-639 (1997)
  MEDLINE   98080636
REFERENCE   2  (bases 1 to 39)
  AUTHORS   Duff,R.J. and Nickrent,D.L.
  TITLE     Mutation rates and phylogenetic utility of mitochondrial SSU rDNA
            in holoparasitic plants
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 39)
  AUTHORS   Duff,R J  and Nickrent,D.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-DEC-1996) Plant Biology, Southern Illinois
            University, Carbondale, IL 62901-6509, USA
FEATURES             Location/Qualifiers
     source          1. .39
                     /organism="Scybalium jamaicense"
                     /organelle="mitochondrion"
                     /db_xref="taxon:48512"
     rRNA            <1. .>39
                     /product="small subunit ribosomal RNA"
BASE COUNT         10 a     5 c    11 g    13 t
ORIGIN

  Query Match           100.0%;  Score 8;  DB 8;  Length 39;
  Best Local Similarity 100.0%;  Pred. No. 9.8e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      10 AACGTTCG 3

RESULT  45
AR035187/c                             PAT       29-SEP-1999
LOCUS       AR035187       40 bp    DNA
DEFINITION  Sequence 48 from patent US 5871730.
ACCESSION   AR035187
VERSION     AR035187.1  GI:5951855
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 40)
  AUTHORS   Brzezinski,R., Dery,C.V. and Beaulieu,C.
  TITLE     Thermostable xylanase DNA, protein and methods of use
  JOURNAL   Patent: US 5871730-A 48 16-FEB-1999;
FEATURES             Location/Qualifiers
     source          1. .40
                     /organism="unknown"
BASE COUNT         10 a    13 c     8 g     9 t
ORIGIN

  Query Match           100.0%;  Score 8;  DB 6;  Length 40;
  Best Local Similarity 100.0%;  Pred. No. 9.8e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      15 AACGTTCG 8


Search completed: November 29, 2001, 14:47:06
Job time: 8319 sec
```

frag1.res

Mon Dec 3 08:02:28 2001

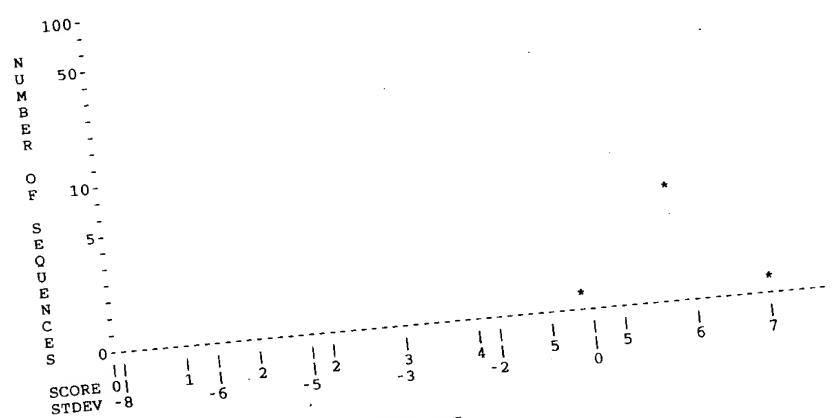
> O <
O| |O IntelliGenetics
> O <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file frag1.res made by sdavid on Wed 28 Nov 101 14:17:43-PST.

Query sequence being compared: FRAG1 (1-8)
Number of sequences searched: 6
Number of scores above cutoff: 6

Results of the initial comparison of FRAG1 (1-8) with:
File : hpvcomplete.seq



PARAMETERS

Similarity matrix	Unitary	K-tuple	4
Mismatch penalty	1	Joining penalty	30
Gap penalty	1.00	Window size	8
Gap size penalty	0.33		
Cutoff score	0		
Randomization group	0		

SEARCH STATISTICS

	Mean	Median	Standard Deviation
Scores:	6	7	0.63

	CPU	Total Elapsed
Times:	00:00:00.00	00:00:00.00

Number of residues: 47269
Number of sequences searched: 6
Number of scores above cutoff: 6

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Init. Opt. Score Score	Sig.	Frame
1. pph16	**** 1 standard deviation above mean **** TOIG of: pph16 check: 6074	7904	7 7	1.58	0
2. hpu06714	**** 0 standard deviation from mean **** TOIG of: hpu06714 check: 486	7801	6 7	0.00	0
3. af131950	TOIG of: af131950 check: 557	7812	6 7	0.00	0
4. a12360	TOIG of: a12360 check: 1580	7909	6 6	0.00	0
5. pph31a	TOIG of: pph31a check: 5866	7912	6 6	0.00	0
6. pph11	**** 1 standard deviation below mean **** TOIG of: pph11 check: 3689	7931	5 5	-1.58	0

1. FRAG1 (1-8)
pph16 TOIG of: pph16 check: 6074 from: 1 to: 7904

TOIG of: pph16 check: 6074 from: 1 to: 7904

LOCUS PPH16 7904 bp DNA circular VRL 18-MAR-1994
DEFINITION Human papillomavirus type 16 (HPV16), complete genome.
ACCESSION K02718
VERSION K02718.1 GI:333031
KEYWORDS circular; complete genome.
SOURCE Papilloma virus type 16 DNA, isolated from a human invasive cervical carcinoma.
ORGANISM Human papillomavirus type 16
Viruses; dsDNA viruses, no RNA stage; Papillomaviridae; Papillomavirus.
REFERENCE 1 (bases 1 to 7904)
AUTHORS Seedorf,K., Kraemmer,G., Duerst,M., Suhai,S. and Roewekamp,W.G.
TITLE Human papillomavirus type 16 DNA sequence
JOURNAL Virology 145, 181-185 (1985)
MEDLINE 85246220
REFERENCE 2 (sites)
AUTHORS Kennedy,I.M., Haddow,J.K. and Clements,J.B.
TITLE A negative element in the human poapillomavirus type 16 genome acts at the level of late mRNA stability
JOURNAL J. Virol. 65, 2093-2097 (1991)
MEDLINE 91162763
COMMENT The sense strand of this double-stranded circular genome is shown, with a numbering system matching the first 60 bp of HPVa1, HPV6b and BPV1. The annotation of sites and features is solely based upon homology comparison with these other papillomaviruses. In addition to the coding sequences reported below, the authors note open reading frames which do not start with 'ATG', but which are found in other papillomaviruses. In particular, a second portion of the E1 gene may be located out to base 2813 (the E1 protein is thought to be generally involved in DNA replication).
A potential 'CAT'-box region is found beginning at base 7895 below, and 'TATA' boxes for early and late transcripts may be located at 17, 65 and 4289. Potential polyadenylation signals are at bases 4213 and 7260.
HPV16, in comparison to HPV types 6 and 11, is more often associated with malignant genital cancers in humans.

FEATURES Location/Qualifiers
source 1. .7904
/organism="Human papillomavirus type 16"
/db_xref="taxon:10581"
TATA_signal 17. .23
TATA_signal 65. .71
gene 83. .559
/gene="E6"
CDS 83. .559
/gene="E6"
/note="E6 ORF from 65 to 559; putative"
/codon_start=1
/product="transforming protein"
/protein_id="AAA46939.1"
/db_xref="GI:333032"
/translation="MHQKRTAMFQDPQERPRKLPQLCTELQTTIHDIILECVYCKQQL
LRREVYDFAFRDLCIVYRDGNPYAVCDKCLKFYSKISEYRHYCYSLYGTTLEQQYNKP
LCDLLIRCINCQKPLCPEEKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL"

```
     gene            562. .858
     CDS             /gene="E7"
                     562. .858
                     /gene="E7"
                     /note="E7 ORF from 544 to 858; putative"
                     /codon_start=1
                     /product="transforming protein"
                     /protein_id="AAA46940.1"
                     /db_xref="GI:333033"
                     /translation="MHGDTPTLHEYMLDLQPETTDLYCYEQLNDSSEEEDEIDGPAGQ
     gene            AEPDRAHYNIVTFCCKCDSTLRLCVQSTHVDIRTLEDLLMGTLGIVCPICSQKP"
     CDS             join(865. .1140,1140. .2813)
                     /gene="E1"
                     join(865. .1140,1140. .2813)
                     /gene="E1"
                     /note="E1 interrupted ORF from 859 to 2813; putative"
                     /codon_start=1
                     /product="replication protein"
                     /protein_id="AAA46936.1"
                     /db_xref="GI:459912"
                     /translation="MADPAGTNGEEGTGCNGWFYVEAVVEKKTGDAISDDENENDSDT
                     GEDLVDFIVNDNDYLTQAETETAHALFTAQEAKQHRDAVQVLKRKYLVSPLSDISGCV
                     DNNISPRLKAICIEKQSRAAKRRLFESEDSGYGNTEVETQQMLQVEGRHETETPCSQY
                     SGGSGGGCSQYSSGSGGEGVSERHTICQTPLTNILNVLKTSNAKAAMLAKFKELYGVS
                     FSELVRPFKSNKSTCCDWCIAAFGLTPSIADSIKTLLQQYCLYLHIQSLACSWGMVVL
                     LLVRYKCGKNRETIEKLLSKLLCVSPMCMMIEPPKLRSTAAALYWYKTGISNISEVYG
                     DTPEWIQRQTVLQHSFNDCTFELSQMVQWAYDNDIVDDSEIAYKYAQLADTNSNASAF
                     LKSNSQAKIVKDCATMCRHYKRAEKKQMSMSQWIKYRCDRVDDGGDWKQIVMFLRYQG
                     VEFMSFLTALKRFLQGIPKKNCILLYGAANTGKSLFGMSLMKFLQGSVICFVNSKSHF
                     WLQPLADAKIGMLDDATVPCWNYIDDNLRNALDGNLVSMDVKHRPLVQLKCPPLLITS
     gene            NTNAGTDSRWPYLHNRLVVFTFPNEFPFDENGNPVYELNDKNWKSFFSRTWSRLSLHE
     CDS             DEDKENDGDSLPTFKCVSGQNTNTL"
                     2755. .3852
                     /gene="E2"
                     2755. .3852
                     /gene="E2"
                     /note="E2 ORF from 2725 to 3852; putative"
                     /codon_start=1
                     /product="regulatory protein"
                     /protein_id="AAA46941.1"
                     /db_xref="GI:333035"
                     /translation="METLCQRLNVCQDKILTHYENDSTDLRDHIDYWKHMRLECAIYY
                     KAREMGFKHINHQVVPTLAVSKNKALQAIELQLTLETIYNSQYSNEKWTLQDVSLEVV
                     LTAPTGCIKKHGYTVEVQFDGDICNTMHYTNWTHIYICEEASVTVVEGQVDYYGLYYV
                     HEGIRTYFVQFKDDAEKYSKNKVWEVHAGGQVILCPTSVFSSNEVSSPEIIRQHLANH
     gene            PAATHTKAVALGTEETQTTIQRPRSEPDTGNPCHTTKLLHRDSVDSAPILTAFNSSHK
     CDS             GRINCNSNTTPIVHLKGDANTLKCLRYRFKKHCTLYTAVSSTWHWTGHNVKHKSAIVT
                     LTYDSEWQRDQFLSQVKIPKTITVSTGFMSI"
                     3332. .3619
                     /gene="E4"
                     <3332. .3619
                     /gene="E4"
                     /note="E4 ORF from 3332 to 3619; putative"
                     /codon_start=1
                     /protein_id="AAA46937.1"
                     /db_xref="GI:459913"
     gene            /translation="YYVLHLCLAATKYPLLKLLGSTWPTTPPRPIPKPSPWAPKKHRR
     CDS             LSSDQDQSQTPETPATPLSCCTETQWTVLQSSLHLTAHTKDGLTVIVTLHP"
                     3863. .4099
                     /gene="E5"
                     <3863. .4099
                     /gene="E5"
                     /note="E5 ORF from 3863 to 4099; putative"
                     /codon_start=1
                     /protein_id="AAA46938.1"
                     /db_xref="GI:459914"
     polyA_signal    /translation="YCIHNITGVLFALLCVLLCVCLLIRPLLLSVSTYTSLIILVLLL
                     WITAASAFRCFIVYIIFVYIPLFLIHTHARFLIT"
     gene            4213. .4218
                     /note="putative"
     CDS             4235. .5656
                     /gene="L2"
                     4235. .5656
                     /gene="L2"
                     /note="L2 ORF from 4133 to 5656; putative"
                     /codon_start=1
                     /product="minor capsid protein"
                     /protein_id="AAA46942.1"
                     /db_xref="GI:333036"
                     /translation="MRHKRSAKRTKRASATQLYKTCKQAGTCPPDIIPKVEGKTIAEQ
                     ILQYGSMGVFFGGLGIGTGSGTGGRTGYIPLGTRPPTATDTLAPVRPPLTVDPVGPSD
                     PSIVSLVEETSFIDAGAPTSVPSIPPDVSGFSITTSTDTTPAILDINNTVTTVTTHNN
                     PTFTDPSVLQPPTPAETGGHFTLSSSTISTHNYEEIPMDTFIVSTNPNTVTSSTPIPG
                     SRPVARLGLYSRTTQQVKVVDPAFVTTPTKLITYDNPAYEGIDVDNTLYFSSNDNSIN
                     IAPDPDFLDIVALHRPALTSRRTGIRYSRIGNKQTLRTRSGKSIGAKVHYYYDLSTID
                     PAEEIELQTITPSTYTTTSHAASPTSINNGLYDIYADDFITDTSTTPVPSVPSTSLSG
                     YIPANTTIPFGGAYNIPLVSGPDIPINITDQAPSLIPIVPGSPQYTIIADAGDFYLHP
     TATA_signal     SYYMLRKRRKRLPYFFSDVSLAA"
                     4289. .4295
     gene            /gene="L2"
                     5559. .7154
     CDS             /gene="L1"
                     5559. .7154
                     /gene="L1"
                     /note="L1 ORF from 5526 to 7154; putative"
                     /codon_start=1
                     /product="major capsid protein"
                     /protein_id="AAA46943.1"
                     /db_xref="GI:333037"
                     /translation="MQVTFIYILVITCYENDVNVYHIFFQMSLWLPSEATVYLPPVPV
                     SKVVSTDEYVARTNIYYHAGTSRLLAVGHPYFPIKKPNNNKILVPKVSGLQYRVFRIH
                     LPDPNKFGFPDTSFYNPDTQRLVWACVGVEVGRGQPLGVGISGHPLLNKLDDTENASA
                     YAANAGVDNRECISMDYKQTQLCLIGCKPPIGEHWGKGSPCTNVAVNPGDCPPLELIN
                     TVIQDGDMVHTGFGAMDFTTLQANKSEVPLDICTSICKYPDYIKMVSEPYGDSLFFYL
                     RREQMFVRHLFNRAGTVGENVPDDLYIKGSGSTANLASSNYFPTPSGSMVTSDAQIFN
                     KPYWLQRAQGHNNGICWGNQLFVTVVDTTRSTNMSLCAAISTSETTYKNTNFKEYLRH
                     GEEYDLQFIFQLCKITLTADVMTYIHSMNSTILEDWNFGLQPPPGGTLEDTYRFVTQA
     polyA_signal    IACQKHTPPAPKEDDPLKKYTFWEVNLKEKFSADLDQFPLGRKFLLQAGLKAKPKFTL
                     GKRKATPTTSSTSTTAKRKKRKL"
                     7260. .7265
BASE COUNT      2601 a    1377 c    1509 g    2417 t
ORIGIN          Unreported.

PPH16  Length: 7904  November 28, 2001 14:10  Type: N  Check: 6074  ..

Initial Score   =      7  Optimized Score  =      7  Significance  =  1.58
Residue Identity =    87% Matches          =      7  Mismatches    =     1
Gaps            =      0  Conservative Substitutions               =     0

                                                                X       X
TTTGTTTGTTTGTTTTTTAATAAACTGTTATTACTTAACAATGCGACACAAACGTTCTGCAAAACGCACAAA
  4200      4210      4220      4230      4240      4250      4260
                                                                AACGTTCG
                                                                ||||||||
ACGTGCATCGGCTACCCAACTTTATAAAACATGCAA
4270      4280      4290      4300

2. FRAG1 (1-8)
   hpu06714        TOIG of: hpu06714  check: 4862  from: 1  to: 7801

 TOIG of: hpu06714  check: 4862  from: 1  to: 7801

LOCUS       HPU06714      7801 bp    DNA                     VRL       04-FEB-1997
DEFINITION  Human papillomavirus HPV-1A (3-3), complete genome.
ACCESSION   U06714
VERSION     U06714.1  GI:458704
KEYWORDS    .
SOURCE      Human papillomavirus.
  ORGANISM  Human papillomavirus
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 7801)
  AUTHORS   Danos,O., Katinka,M. and Yaniv,M.
  TITLE     Human papillomavirus 1A complete DNA sequence a novel type of
```

```
            genome organization among Papovaviridae
  JOURNAL   EMBO J. 1, 231-236 (1982)
  MEDLINE   84182467
REFERENCE   2  (bases 1 to 7801)
  AUTHORS   Meissner,J.
  TITLE     Complete nucleotide sequencing of an HPV-1a variant and
            determination of extant errors in the prototype HPV-1a sequence
  JOURNAL   Virus Genes 9 (2), 189-191 (1995)
  MEDLINE   95250312
REFERENCE   3  (bases 1 to 7801)
  AUTHORS   Meissner,J.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-FEB-1994) John D. Meissner, Duke University Medical
            Center, Microbiology, 277 Carl Building, Research Drive, Durham, NC
            27710 USA
FEATURES             Location/Qualifiers
     source          1. .7801
                     /organism="Human papillomavirus"
                     /strain="HPV-1A (3-3)"
                     /db_xref="taxon:10566"
     variation       142
                     /replace="a"
     conflict        1283
                     /citation=[1]
                     /replace="a"
     conflict        2301
                     /citation=[1]
                     /replace="t"
     variation       2825
                     /replace="t"
     conflict        3884. .3886
                     /citation=[1]
                     /replace="aa"
     conflict        4332
                     /citation=[1]
                     /replace="a"
     conflict        4376. .4382
                     /citation=[1]
                     /replace="gagggaa"
     variation       5794
                     /replace="a"
     variation       6905
                     /replace="g"
     variation       7186. .7187
                     /note="15 bp deletion"
                     /citation=[1]
                     /replace="ctagattgtattgctat"
     variation       7560
                     /replace="c"
     variation       7618
                     /replace="g"
     conflict        7677. .7678
                     /citation=[1]
                     /replace="cc"
     variation       7787
                     /replace="g"
BASE COUNT     2389 a    1482 c    1664 g    2266 t
ORIGIN

HPU06714  Length: 7801  November 28, 2001 14:10  Type: N  Check: 4862  ..

Initial Score  =      6  Optimized Score  =      7  Significance  =   0.00
Residue Identity =   87%  Matches         =      7  Mismatches    =      1
Gaps           =      0  Conservative Substitutions               =      0

                                                  X       X
                                                  AACGTTCG
                                                  |||||| |
AAAGGTACTGAAAACGATTGGTTTTTGGTGGAGGCGACAGATTGTGAGGAAACGTTAGAGGAAACCTCACTT
    830       840       850       860       870   X   880       890

GGTGACCTAGATAATGTTTCTTGTGTTAGCGACTTA
  900       910       920       930
```

```
3. FRAG1 (1-8)
   af131950     TOIG of: af131950  check: 557  from: 1  to: 7812

 TOIG of: af131950  check: 557  from: 1  to: 7812

LOCUS       AF131950     7812 bp    DNA              VRL       07-FEB-2001
DEFINITION  Human papillomavirus candHPV85, complete genome.
ACCESSION   AF131950
VERSION     AF131950.1  GI:4574720
KEYWORDS    .
SOURCE      Human papillomavirus candHPV85
  ORGANISM  Human papillomavirus candHPV85
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 7812)
  AUTHORS   Chow,V.T.K. and Leong,P.W.F.
  TITLE     Complete nucleotide sequence, genomic organization and phylogenetic
            analysis of a novel genital human papillomavirus type, HLT7474-S
  JOURNAL   J. Gen. Virol. 80 (Pt 11), 2923-2929 (1999)
  MEDLINE   20047972
   PUBMED   10580054
REFERENCE   2  (bases 1 to 7812)
  AUTHORS   Chow,V.T.K. and Leong,W.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-FEB-1999) Department of Microbiology, National
            University of Singapore, 10 Kent Ridge Crescent, Singapore 119260,
            Singapore
FEATURES             Location/Qualifiers
     source          1. .7812
                     /organism="Human papillomavirus candHPV85"
                     /db_xref="taxon:151757"
                     /note="isolated from scraped uterine cervical cells from a
                     female sex worker; overlapping PCR products"
     protein_bind    40. .51
                     /note="putative"
                     /bound_moiety="E2 protein"
                     /function="transcriptional regulation"
     protein_bind    56. .67
                     /note="putative"
                     /bound_moiety="E2 protein"
                     /function="transcriptional regulation"
     TATA_signal     71. .79
                     /note="putative"
     gene            105. .578
                     /gene="E6"
     CDS             105. .578
                     /gene="E6"
                     /codon_start=1
                     /product="putative transforming protein E6"
                     /protein_id="AAD24181.1"
                     /db_xref="GI:4574721"
                     /translation="MAEFGNPATRPYKLPDLCNTLDTSLQDIEISCVYCKSVLQRTEV
                     YEFAFADLFVVYRDGIPYAACQNCLMFYSKIRELRYYSDSVYGETLEKLTNSNIYDIL
                     IRCLRCQKPLCPAEKLKHLNEKRRFHKIAGKYRGQCRRCMTRAQEQQGSRRETQV"
     gene            587. .913
                     /gene="E7"
     CDS             587. .913
                     /gene="E7"
                     /codon_start=1
                     /product="putative transforming protein E7"
                     /protein_id="AAD24182.1"
                     /db_xref="GI:4574722"
                     /translation="MHGPKPTVHDIVLDLEPYNEVQEVDLYCYEELNNSEEEIDEPDN
                     AINHRQPLLARREELQRHTICCVCCKCEASLQLVVESSAADLRDLQQLFLGTLSFLCP
                     LCAVLR"
     gene            920. .2866
                     /gene="E1"
     CDS             920. .2866
                     /gene="E1"
                     /codon_start=1
                     /product="putative replication protein E1"
```

```
                     /protein_id="AAD24183.1"
                     /db_xref="GI:4574723"
                     /translation="MADPEGTDGDGTGCNGWFFVQAIVDKKTGDTISEDEDEDGTDTG
                     SDLVDFIDTSNTNYMQADREAAQALLHAQEVETDTKLLHALKRKYGAHSTENSPCRDT
                     ASIHSNLSSPLQEISLNSYNNTAKRRLCSVPDSGYGNTQVETLQTQVTLDTNVFGDGK
                     NGDGLNSEACSTDNEMDIENQNPNSPMTQIVSLLKVNNKKAAILAKFKETYGLSFTDL
                     VRTFKSDKTTCTDWVAAICGVNPNIAEGFKTLIQPYVLYAHIQCMDCSWGVFILALLR
                     YKCGKNRLTVAKGLSTLLHVPDTHMLIEPPKLRSSCAALYWYRTGISNISEVTGDTPE
                     WIQRQTIIQHGIDDSVFDLSEMIQWAFDNDYIDESDIAYEYAQLADCNSNAAAFLKSN
                     CQAKYLRDCAVMCRHYKRAQRKQMNMSQWISYRCDKIDDGGDWKPIVQFLRFQGIEFI
                     TFLRAFKDFLKGTPKKNCIVIYGPANTGKSYFCMSLIQFLHGTVLSFVNSNSHFWLEP
                     LTDTKIAMVDDATPTCWSYFDNYMRNALDGNPISIDRKHKHLIQMKCPPMLITSNTNP
                     ATDDRWPYLRSRVTVFTFPHTFPFDSNGNPVYDINDKNWKCFFKRTWSRLDLHQEEED
gene                 TENDGCPIPTFKCVTGKDSRTLRD"
                     2799. .3932
CDS                  /gene="E2"
                     2799. .3932
                     /gene="E2"
                     /codon_start=1
                     /product="putative regulatory protein E2"
                     /protein_id="AAD24184.1"
                     /db_xref="GI:4574724"
                     /translation="MMDALSQRLNVLQEKILEHYETDSTDICDQIDYWKCVRLENAIY
                     YAAREHGLKTINHHVVPTFQISKSKAHEAIELQMALESLAKSEFKNELWTLQDTCQEL
                     YQTPPQQCFKKQGQTVEVRYDGDKDNTMHYTSWDYIYYVTEGDKWCKTKGYVNYCGLY
                     YIKEGQQTYYVQFKCDAQQYGQSGKWEVWYNGKKIECVNTCDKIIQCSESMYSTCDET
                     VSATAIARELQHPTTPYTEATTVCTQKSGGSAPTRNPFRHCGFTETSEVDGLSVDHLN
gene                 NPLLSAPAGNNGFRKHNSGDTTPIVHLKGDKNRLKCLRYRLQKFHNLYTQISCTWHWI
                     GGKETTRTGILTLAYANETQRQKFLDVVKIPNTVNVSLGYMTL"
CDS                  <3412. .3699
                     /gene="E4"
                     <3412. .3699
                     /gene="E4"
                     /note="start codon absent"
                     /codon_start=1
                     /product="putative protein E4"
                     /protein_id="AAD24185.1"
                     /db_xref="GI:4574725"
                     /translation="YNVLNLCTVPVTRQYPLLQLLENYNTPPHRIPKPPPCAPKRAGV
gene                 RRRLETHSDIVDSQKQVKSTDCPWTTSTTPCLVHLQATTGSGSTIVVTLRL"
                     3940. .4161
CDS                  /gene="E5"
                     3940. .4161
                     /gene="E5"
                     /codon_start=1
                     /product="putative protein E5"
                     /protein_id="AAD24186.1"
                     /db_xref="GI:4574726"
                     /translation="MLIAVFLLCFCVCLYVCCTIPLFPATNVCAHAWLLVFLLIVFIT
gene                 TPFTAFIVYIFFFILPMFLLHLHSVHTFD"
                     4226. .5629
CDS                  /gene="L2"
                     4226. .5629
                     /gene="L2"
                     /codon_start=1
                     /product="putative minor capsid protein L2"
                     /protein_id="AAD24187.1"
                     /db_xref="GI:4574727"
                     /translation="MVSHRAARRKRASATDLYKTCKQSGTCPPDVINKVEGTTLADKL
                     LQWTSLGIFLGGLGIGTGSGTGGRTGYIPLGGRPNTVVDVSPARPPVVIESVGPSDPS
                     IVTLVEESSIVTSGAPVPTFTGTSGFEITSSATTTPAVLDITPASGSVQLSSTSFTNP
                     AFTDPSVIEVPQTGEVSGDIFITTPTSGTHGYEEIPMHTFATQGRGTEPISSTPIPGV
                     RRVAGPRLYSQAYQQVKITNSDFISRPSTLVTFTNPAYEPIDTTLTFSPQDVVPDPDF
                     MDIVRLHRPALTSRRGTVRFSRLGKKLTMSTRSGKQIGAQVHYYHDISPISHIGESIE
                     MQPLLPDAAVTADTNGLFDIYADTDIDNNAMLYDRNISDVTQPTTSTISSVSSRYSNT
gene                 TIPLATSWDVPVHTGPDMTLPTTIPQWPNIVPLLPNNTHSVVLQGTNYYLWPNYYFIF
                     KKRKRVPYFLTDGFVAF"
CDS                  5610. .7124
                     /gene="L1"
                     5610. .7124
                     /gene="L1"
                     /codon_start=1
                     /product="putative major capsid protein L1"
                     /protein_id="AAD24188.1"
                     /db_xref="GI:4574728"
                     /translation="MALWRSSDSKVYLPPPSVAKVVNTDDYVTRTSTFYHAGSSRLLT
                     VGHPYYKVTSNGGRKQDIPKVSAYQYRVFRVTLPDPNKFGLPDTNIYNPETQRLVWAC
                     VGMEVGRGQPLGVSLSGHPFYNKLDDTENSHVATSVVTHDTRDNVSVDYKQTQLCIIG
                     CVPAIGEHWAKGTACKPGAVQTGDCPPLELVNTPIEDGDMIDTGYGAMDFSTLQDNKS
                     EVPLDICQSICKYPDYLQMSADAYGDSMFFCLRREQLFARHFWNRGGTIGDAVPETLY
                     IKGTNDRATPGSCIYSPSPSGSMVSSDAQMFNKPYWLHKAQGQNNGICWHNQLFITVV
                     DTTRSTNLTLSTATTNPVPSIYEPSKFKEYTRHVEEYDLQFIFQLCKITLTTDVMSYI
                     HNMDPTILDSWNFGVSPPPSASLVDTYRFLQSSAITCQKDVVVPQKKDPYEKLKFWNV
                     DLKEHFSSDLDQFPLGRKFLLQAGLRPKPTIGPRKRVASTSTATRPSKRKRTAK"
protein_bind         7456. .7467
                     /note="putative"
                     /bound_moiety="E2 protein"
misc_feature         /function="transcriptional regulation"
                     7647. .7666
protein_bind         /note="consensus motif of oncogenic anogenital HPV types"
                     7776. .7787
                     /note="putative"
                     /bound_moiety="E2 protein"
                     /function="transcriptional regulation"
BASE COUNT       2502 a    1391 c    1558 g    2361 t
ORIGIN
```

AF131950 Length: 7812 November 28, 2001 14:10 Type: N Check: 557 ..

Initial Score = 6 Optimized Score = 7 Significance = 0.00
Residue Identity = 87% Matches = 7 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

```
                                                                X       X
                                                                AACGTTCG
                                                                |||||| |
  TATTCTAAAATAAGAGAACTAAGATATTATAGCGACTCTGTGTATGGGGAAACGTTAGAAAAACTAACAAAT
320         330         340         350         360         370     X   380
  AGCAATATATATGATATATTAATAAGGTGTTTACGG
390         400         410         420
```

4. FRAG1 (1-8)
a12360 TOIG of: a12360 check: 1580 from: 1 to: 7909

TOIG of: a12360 check: 1580 from: 1 to: 7909

```
LOCUS       A12360          7909 bp    DNA                     PAT       12-DEC-1993
DEFINITION  complete nucleotide sequence of HPV-33.
ACCESSION   A12360
VERSION     A12360.1  GI:492936
KEYWORDS    .
SOURCE      Human papillomavirus type 33.
  ORGANISM  Human papillomavirus type 33
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 7909)
  AUTHORS   .
  JOURNAL   Patent: WO 8705630-A 1 24-SEP-1987;
FEATURES             Location/Qualifiers
     source          1. .7909
                     /organism="Human papillomavirus type 33"
                     /db_xref="taxon:10586"
BASE COUNT       2544 a    1354 c    1535 g    2474 t      2 others
ORIGIN
```

A12360 Length: 7909 November 28, 2001 14:10 Type: N Check: 1580 ..

Initial Score = 6 Optimized Score = 6 Significance = 0.00
Residue Identity = 75% Matches = 6 Mismatches = 2
Gaps = 0 Conservative Substitutions = 0

```
                                                                X       X
                                                                AACGTTCG
                                                                ||||||
```

```
   TAGAGAAACTGCACTGTGACGTGTAAAAACGCCATGAGAGGACACAAGCCAACGTTAAAGGAATATGTTTTA
540        550        560        570        580        590        600        610

   GATTTATATCCTGAACCAACTGACCTATACTGCTAT
        620        630        640


5  FRAG1 (1-8)
   pph31a        TOIG of: pph31a  check: 5866  from: 1  to: 7912

   TOIG of: pph31a  check: 5866  from: 1  to: 7912

LOCUS       PPH31A       7912 bp    DNA   circular  VRL       18-MAR-1994
DEFINITION  Human papillomavirus type 31 (HPV-31) complete genome.
ACCESSION   J04353
VERSION     J04353.1  GI:333048
KEYWORDS    complete genome.
SOURCE      Human papillomavirus type 31
  ORGANISM  Human papillomavirus type 31
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 7912)
  AUTHORS   Goldsborough,M.D., DiSelvestre,D., Temple,G.F. and Lorincz,A.T.
  TITLE     Nucleotide sequence of human papillomavirus type 31: A cervical
            neoplasia associated virus
  JOURNAL   Virology 171, 306-311 (1989)
  MEDLINE   89299478
COMMENT     Draft entry and computer-readable copy of sequence [1] kindly
            submitted by M.D.Goldsborough, 05-JUL-1989.
FEATURES             Location/Qualifiers
     source          1. .7912
                     /organism="Human papillomavirus type 31"
                     /db_xref="taxon:10585"
     TATA_signal     19. .24
     TATA_signal     69. .74
     gene            108. .557
                     /gene="E6"
     CDS             108. .557
                     /gene="E6"
                     /note="ORF E6 from bp 39 to 557"
                     /codon_start=1
                     /product="transforming protein"
                     /protein_id="AAA46950.1"
                     /db_xref="GI:459916"
                     /translation="MFKNPAERPRKLHELSSALEIPYDELRLNCVYCKGQLTETEVLD
                     FAFTDLTIVYRDDTPHGVCTKCLRFYSKVSEFRWYRYSVYGTTLEKLTNKGICDLLIR
                     CITCQRPLCPEEKQRHLDKKKRFHNIGGRWTGRCIACWRRPRTETQV"
     misc_feature    228. .236
                     /gene="E6"
                     /standard_name="Splice donor"
     misc_feature    403. .414
                     /gene="E6"
                     /standard_name="Splice acceptor"
     gene            560. .856
                     /gene="E7"
     CDS             560. .856
                     /gene="E7"
                     /note="ORF E7 from bp 545 to 856"
                     /codon_start=1
                     /product="transforming protein"
                     /protein_id="AAA46951.1"
                     /db_xref="GI:459917"
                     /translation="MRGETPTLQDYVLDLQPEATDLHCYEQLPDSSDEEDVIDSPAGQ
                     AEPDTSNYNIVTFCCQCKSTLRLCVQSTQVDIRILQELLMGSFGIVCPNCSTRL"
     gene            862. .2751
                     /gene="E1"
     CDS             862. .2751
                     /gene="E1"
                     /note="ORF E1 from bp 850 to bp 2751"
                     /codon_start=1
                     /product="replication protein"
                     /protein_id="AAA46952.1"
                     /db_xref="GI 459918"
                     /translation="MADPAGTDGEGTGCNGWFYVEAVIDRQTGDNISEDENEDSSDTG
                     EDMVDFIDNCNVYNNQAEAETAQALFHAQEAEEHAEAVQVLKRKYVGSPLSDISSCVD
                     YNISPRLKAICIENNSKTAKRRLFELPDSGYGNTEVETQQMVQVEEQQTTLSCNGSDG
                     THSERENETPTRNILQVLKTSNGKAAMLGKFKELYGVSFMELIRPFQSNKSTCTDWCV
                     AAFGVTGTVAEGFKTLLQPYCLYCHLQSLACSWGMVMLMLVRFKCAKNRITIEKLLEK
                     LLCISTNCMLIQPPKLRSTAAALYWYRTGMSNISDVYGETPEWIERQTVLQHSFNDTT
                     FDLSQMVQWAYDNDVMDDSEIAYKYAQLADSDSNACAFLKSNSQAKIVKDCGTMCRHY
                     KRAEKRQMSMGQWIKSRCDKVSDEGDWRDIVKFLRYQQIEFVSFLSALKLFLKGVPKK
                     NCILIHGAPNTGKSYFGMSLISFLQGCIISYANSKSHFWLQPLADAKIGMLDDATTPC
                     WHYIDNYLRNALDGNPVSIDVKHKALMQLKCPPLLITSNINAGKDDRWPYLHSRLVVF
                     TFPNPFPFDKNGNPVYELSDKNWKSFFSRTWCRLNLHEEEDKENDGDSFSTFKCVSGQ
                     NIRTL"
     gene            2693. .3811
                     /gene="E2"
     CDS             2693. .3811
                     /gene="E2"
                     /note="ORF E2 from bp 2663 to 3811"
                     /codon_start=1
                     /product="regulatory protein"
                     /protein_id="AAA46953.1"
                     /db_xref="GI:459919"
                     /translation="METLSQRLNVCQDKILEHYENDSKRLCDHIDYWKHIRLECVLMY
                     KAREMGIHSINHQVVPALSVSKAKALQAIELQMMLETLNNTEYKNEDWTMQQTSLELY
                     LTAPTGCLKKHGYTVEVQFDGDVHNTMHYTNWKFIYLCIDGQCTVVEGQVNCKGIYYV
                     HEGHITYFVNFTEEAKKYGTGKKWEVHAGGQVIVFPESVFSSDEISFAGIVTKLPTAN
                     NTTTSNSKTCALGTSEGVRRATTSTKRPRTEPEHRNTHHPNKLLRGDSVDSVNCGVIS
                     AAACTNQTRAVSCPATTPIIHLKGDANILKCLRYRLSKYKQLYEQVSSTWHWTCTDGK
                     HKNAIVTLTYISTSQRDDFLNTVKIPNTVSVSTGYMTI"
     gene            3270. .3578
                     /gene="E4"
     CDS             <3270. .3578
                     /gene="E4"
                     /note="ORF E4 from bp 3270 to 3578"
                     /codon_start=1
                     /protein_id="AAA46949.1"
                     /db_xref="GI:459915"
                     /translation="LFFLNLYLAVTKYPLLGLLQSYQQPTTPPHRIPKPAPWAPVKVC
                     GGRRRLLSDQEQSQSTETPTTPTSCCEATPWTVSTVGLSVQLHAQTKQGLSVVLQLHL
                     "
     gene            3816. .4070
                     /gene="E5"
     CDS             3816. .4070
                     /gene="E5"
                     /note="ORF E5 from bp 3804 to 4070"
                     /codon_start=1
                     /protein_id="AAA46954.1"
                     /db_xref="GI:459920"
                     /translation="MIELNISTVSIVLCFLLCFCVLLFVCLVIRPLVLSVSVYATLLL
                     LIVILWVIATSPLRCFCIYVVFIYIPLFVIHTHASFLSQQ"
     repeat_region   4099. .4134
     polyA_signal    4138. .4143
                     /note="putative"
     repeat_region   4143. .4158
     gene            4171. .5571
                     /gene="L2"
     CDS             4171. .5571
                     /gene="L2"
                     /note="ORF L2 from bp 4060 to bp 5571"
                     /codon_start=1
                     /product="minor capsid protein"
                     /protein_id="AAA46955.1"
                     /db_xref="GI:459921"
                     /translation="MRSKRSTKRTKRASATQLYQTCKAAGTCPSDVIPKIEHTTIADQ
                     ILRYGSMGVFFGGLGIGSGSGTGGRTGYVPLSTRPSTVSEASIPIRPPVSIDPVGPLD
                     PSIVSLVEESGIVDVGAPAPIPHPPTTSGFDIATTADTTPAILDVTSVSTHENPTFTD
                     PSVLQPPTPAETSGHLLLSSSSISTHNYEEIPMDTFIVSTNNENITSSTPIPGVRRPA
                     RLGLYSKATQQVKVIDPTFLSAPKQLITYENPAYETVNAEESLYFSNTSHNIAPDPDF
                     LDIIALHRPALTSRRNTVRYSRLGNKQTLRTRSGATIGARVHYYYDISSINPAGESIE
                     MQPLGASATTTSTLNDGLYDIYADTDFTVDTPATHNVSPSTAVQSTSAVSAYVPTNTT
                     VPLSTGFDIPIFSGPDVPIEHAPTQVFPFPLAPTTPQVSIFVDGGDFYLHPSYYMLKR
                     RRKRVSYFFTDVSVAA"
     gene            5552. .7066
                     /gene="L1"
```

```
     CDS             5552. .7066
                     /gene="L1"
                     /note="ORF L1 from bp 5516 to  7066"
                     /codon_start=1
                     /product="major capsid protein"
                     /protein_id="AAA46956.1"
                     /db_xref="GI:459922"
                     /translation="MSLWRPSEATVYLPPVPVSKVVSTDEYVTRTNIYYHAGSARLLT
                     VGHPYYSIPKSDNPKKIVVPKVSGLQYRVFRVRLPDPNKFGFPDTSFYNPETQRLVWA
                     CVGLEVGRGQPLGVGISGHPLLNKFDDTENSNRYAGGPGTDNRECISMDYKQTQLCLL
                     GCKPPIGEHWGKGSPCSNNAITPGDCPPLELKNSVIQDGDMVDTGFGAMDFTALQDTK
                     SNVPLDICNSICKYPDYLKMVAEPYGDTLFFYLRREQMFVRHFFNRSGTVGESVPTDL
                     YIKGSGSTATLANSTYFPTPSGSMVTSDAQIFNKPYWMQRAQGHNNGICWGNQLFVTV
                     VDTTRSTNMSVCAAIANSDTTFKSSNFKEYLRHGEEFDLQFIFQLCKITLSADIMTYI
                     HSMNPAILEDWNFGLTTPPSGSLEDTYRFVTSQAITCQKTAPQKPKEDPFKDYVFWEV
                     NLKEKFSADLDQFPLGRKFLLQAGYRARPKFKAGKRSAPSASTTTPAKRKKTKK"
     polyA_signal    7227. .7231
     repeat_region   7291. .7302
                     /note="putative"
     repeat_region   7314. .7333
     protein_bind    7406. .7420
                     /note="putative"
                     /standard_name="glucocorticoid responsive element"
     protein_bind    7477. .7488
                     /bound_moiety="hormone receptor"
                     /function="gene transcription"
     enhancer        7542. .7549
                     /bound_moiety="E2"
     protein_bind    7868. .7879
                     /standard_name="keratinocyte-dependant enhancer"
                     /function="gene transcription"
                     /bound_moiety="E2"
BASE COUNT      2528 a    1364 c    1572 g    2448 t
ORIGIN

PPH31A  Length: 7912  November 28, 2001 14:10  Type: N  Check: 5866  ..

Initial Score  =      6  Optimized Score  =      6  Significance  =  0.00
Residue Identity  =  75%  Matches  =      6  Mismatches  =      2
Gaps  =      0  Conservative Substitutions  =      0

                                                    X       X
  CGTTTAGGGTTATATAGTAAGGCTACACAACAAGTAAAAGTTATTGATCCAACGTTTCTTAGTGCTCCAAAA
                                                    AACGTTCG
    4830      4840      4850      4860      4870    X 4880 X    4890

  CAGCTAATTACATATGAAAACCCTGCCTATGAAACT
  4900      4910      4920      4930

6. FRAG1 (1-8)
   pph11          TOIG of: pph11  check: 3689  from: 1  to: 7931

TOIG of: pph11  check: 3689  from: 1  to: 7931

LOCUS       PPH11          7931 bp    DNA    circular  VRL       02-JUN-1994
DEFINITION  Human papillomavirus type 11 (HPV-11) complete genome.
ACCESSION   M14119
VERSION     M14119.1  GI:333026
KEYWORDS    complete genome.
SOURCE      Human laryngeal papillomavirus type 11 DNA.
  ORGANISM  Human papillomavirus type 11
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 7931)
  AUTHORS   Dartmann,K., Schwarz,E., Gissmann,L. and Zur Hausen,H.
  TITLE     The nucleotide sequence and genome organization of human papilloma
            virus type 11
  JOURNAL   Virology 151, 124-130 (1986)
  MEDLINE   86181601
COMMENT     ORF L1 is assumed to encode the major structural protein.
FEATURES             Location/Qualifiers
```

```
     source          1. .7931
                     /organism="Human papillomavirus type 11"
                     /db_xref="taxon:10580"
     CAAT_signal     9. .15
     protein_bind    35. .46
                     /note="putative"
                     /note="putative"
                     /function="gene transcription"
     protein_bind    50. .61
                     /bound_moiety="E2"
                     /note="putative"
                     /function="gene transcription"
     TATA_signal     66. .71
                     /bound_moiety="E2"
     gene            102. .554
                     /note="putative"
     CDS             102. .554
                     /gene="E6"
                     /gene="E6"
                     /note="102 is position of first start codon in ORF E6;
                     putative"
                     /codon_start=1
                     /product="transforming protein"
                     /protein_id="AAA46927.1"
                     /db_xref="GI:496193"
                     /translation="MESKDASTSATSIDQLCKTFNLSLHTLQIQCVFCRNALTTAEIY
                     AYAYKNLKVVWRDNFPFAACACCLELQGKINQYRHFNYAAYAPTVEEETNEDILKVLI
                     RCYLCHKPLCEIEKLKHILGKARFIKLNNQWKGRCLHCWTTCMEDLLP"
     gene            530. .826
     CDS             530. .826
                     /gene="E7"
                     /gene="E7"
                     /note="530 is position of first start codon in ORF E7;
                     putative"
                     /codon_start=1
                     /product="transforming protein"
                     /protein_id="AAA46928.1"
                     /db_xref="GI:496194"
                     /translation="MHGRLVTLKDIVLDLQPPDPVGLHCYEQLEDSSEDEVDKVDKQD
     gene            AQPLTQHYQILTCCCGCDSNVRLVVECTDGDIRQLQDLLLGTLNIVCPICAPKP"
     CDS             832. .2781
                     /gene="E1"
                     832. .2781
                     /gene="E1"
                     /note="832 is position of first start codon in ORF E1;
                     putative"
                     /codon_start=1
                     /product="replication protein"
                     /protein_id="AAA46929.1"
                     /db_xref="GI:496195"
                     /translation="MADDSGTENEGSGCTGWFMVEAIVEHTTGTQISEDEEEEVEDSG
                     YDMVDFIDDRHITQNSVEAQALFNRQEADAHYATVQDLKRKYLGSPYVSPISNVANAV
                     ESEISPRLDAIKLTTQPKKVKRRLFETRELTDSGYGYSEVEAATQVEKHGDPENGGDG
                     QERDTGRDIEGEGVEHREAEAVDDSTREHADTSGILELLKCKDIRSTLHGKFKDCFGL
                     SFVDLIRPFKSDRTTCADWVVAGFGIHHSIADAFQKLIEPLSLYAHIQWLTNAWGMVL
                     LVLIRFKVNKSRCTVARTLGTLLNIPENHMLIEPPKIQSGVRALYWFRTGISNASTVI
                     GEAPEWITRQTVIEHSLADSQFKLTEMVQWAYDNDICEESEIAFEYAQRGDFDSNARA
     gene            FLNSNMQAKYVKDCAIMCRHYKHAEMKKMSIKQWIKYRGTKVDSVGNWKPIVQFLRHQ
     CDS             NIEFIPFLSKLKLWLHGTPKKNCIAIVGPPDTGKSCFCMSLIKFLGGTVISYVNSCSH
                     FWLQPLTDAKVALLDDATQPCWTYMDTYMRNLLDGNPMSIDRKHRALTLIKCPPLLVT
                     SNIDISKEEKYKYLHSRVTTFTFPNPFPFDRNGNAVYELSDANWKCFFERLSSSLDIE
     gene            DSEDEEDGSNSQAFRCVPGSVVRTL"
     CDS             2723. .3826
                     /gene="E2"
                     2723. .3826
                     /gene="E2"
                     /note="2723 is position of first start codon in ORF E2;
                     putative"
                     /codon_start=1
                     /product="regulatory protein"
                     /protein_id="AAA46930.1"
                     /db_xref="GI:496196"
                     /translation="MEAIAKRLDACQDQLLELYEENSIDIHKHIMHWKCIRLESVLLH
```

```
                KAKQMGLSHIGLQVVPPLTVSETKGHNAIEMQMHLESLAKTQYGVEPWTLQDTSYEMW
                LTPPKRCFKKQGNTVEVKFDGCEDNVMEYVVWTHIYLQDNDSWVKVTSSVDAKGIYYT
                CGQFKTYYVNFNKEAQKYGSTNHWEVCYGSTVICSPASVSSTVREVSIAEPTTYTPAQ
                TTAPTVSACTTEDGVSAPPRKRARGPSTNNTLCVANIRSVDSTINNIVTDNYNKHQRR
                NNCHSAATPIVQLQGDSNCLKCFRYRLNDKYKHLFELASSTWHWASPEAPHKNAIVTL
                TYSSEEQRQQFLNSVKIPPTIRHKVGFMSLHLL"
gene            3255. .3581
                /gene="E4"
CDS             3255. .3581
                /gene="E4"
                /note="3255 is position of first start codon in ORF E4;
                putative"
                /codon_start=1
                /protein_id="AAA46931.1"
                /db_xref="GI:496197"
                /translation="MVVPIIGKYVMAAQLYVLLHLYLALYEKYPLLNLLHTPPHRPPP
                LQCPPAPRKTACRRRLGSEHVDRPLTTPCVWPTSDPWTVQSTTSSLTITTSTKEGTTV
                TVQLRL"
gene            3871. .4146
                /gene="E5A"
CDS             3871. .4146
                /gene="E5A"
                /note="3871 is position of first start codon in ORFE5A;
                putative"
                /codon_start=1
                /protein_id="AAA46932.1"
                /db_xref="GI:496198"
                /translation="MEVVPVQIAAATTTTLILPVVIAFAVCILSIVLIILISDFVVYT
                SVLVLTLLLYLLLWLLLTTPLQFFLLTLCVCYFPAFYIHIYIVQTQQ"
gene            4146. .4370
                /gene="E5B"
CDS             4146. .4370
                /gene="E5B"
                /note="4146 is position of first start codon in ORF E5B;
                putative"
                /codon_start=1
                /protein_id="AAA46933.1"
                /db_xref="GI:496199"
                /translation="MVMLTCHLNDGDTWLFLWLFTAFVVAVLGLLLLHYRAVHGTEKT
                KCAKCKSNRNTTVDYVYMSHGDNGDYVYMN"
polyA_signal    4371. .4376
                /note="putative"
gene            4417. .5784
                /gene="L2"
CDS             4417. .5784
                /gene="L2"
                /note="4417 is position of first start codon in ORF L2;
                putative"
                /codon_start=1
                /product="minor capsid protein"
                /protein_id="AAA46934.1"
                /db_xref="GI:496200"
                /translation="MKPRARRRKRASATQLYQTCKATGTCPPDVIPKVEHTTIADQIL
                KWGSLGVFFGGLGIGTGAGSGGRAGYIPLGSSPKPAITGGPAARPPVLVEPVAPSDPS
                IVSLIEESAIINAGAPEVVPPTQGGFTITSSESTTPAILDVSVTNHTTTSVFQNPLFT
                EPSVIQPQPPVEASGHILISAPTITSQHVEDIPLDTFVVSSSDSGPTSSTPLPRAFPR
                PRVGLYSRALQQVQVTDPAFLSTPQRLVTYDNPVYEGEDVSLQFTHESIHNAPDEAFM
                DIIRLHRPAITSRRGLVRFSRIGQRGSMYTRSGQHIGARIHYFQDISPVTQAAEEIEL
                HPLVAAENDTFDIYAEPFDPIPDPVQHSVTQSYLTSTPNTLSQSWGNTTVPLSIPSDW
                FVQSGPDITFPTASMGTPFSPVTPALPTGPVFITGSDFYLHPTWYFARRRRKRIPLFF
                TDVAA"
polyA_signal    4545. .4551
                /gene="L2"
                /note="putative"
gene            5771. .7276
                /gene="L1"
CDS             5771. .7276
                /gene="L1"
                /note="5771 is position of first start codon in ORF L1;
                putative"
                /codon_start=1
                /product="major capsid protein"
                /protein_id="AAA46935.1"
                /db_xref="GI:496201"
                /translation="MWRPSDSTVYVPPPNPVSKVVATDAYVKRTNIFYHASSSRLLAV
                GHPYYSIKKVNKTVVPKVSGYQYRVFKVVLPDPNKFALPDSSLFDPTTQRLVWACTGL
                EVGRGQPLGVGVSGHPLLNKYDDVENSGGYGGNPGQDNRVNVGMDYKQTQLCMVGCAP
                PLGEHWGKGTQCSNTSVQNGDCPPLELITSVIQDGDMVDTGFGAMNFADLQTNKSDVP
                LDICGTVCKYPDYLQMAADPYGDRLFFYLRKEQMFARHFFNRAGTVGEPVPDDLLVKG
                GNNRSSVASSIYVHTPSGSLVSSEAQLFNKPYWLQKAQGHNNGICWGNHLFVTVVDTT
                RSTNMTLCASVVSKSATYTNSDYKEYMRHVEEFDLQFIFQLCSITLSAEVMAYIHTMNP
                SVLEDWNFGLSPPPNGTLEDTYRYVQSQAITCQKPTPEKEKQDPYKDMSFWEVNLKEK
                FSSELDQFPLGRKFLLQSGYRGRTSARTGIKRPAVSKPSTAPKRKRTKTKK"
5'UTR           join(7277. .7931,1. .101)
                /standard_name="LCR"
                /function="regulatory region"
repeat_region   7320. .7360
                /standard_name="direct repeats"
                /note="putative"
repeat_region   7339. .7374
                /standard_name="direct repeats"
                /note="putative"
repeat_region   7374. .7403
                /note="putative"
polyA_signal    7457. .7462
                /note="putative"
protein_bind    7592. .7603
                /note="putative"
                /function="gene transcription"
                /bound_moiety="E2"
polyA_signal    7748. .7753
                /note="putative"
protein_bind    7890. .7901
                /note="putative"
                /function="gene transcription"
                /bound_moiety="E2"
BASE COUNT      2406 a   1519 c   1736 g   2270 t
ORIGIN      4557 bp upstream of HindIII site.

  PPH11  Length: 7931  November 28, 2001 14:10  Type: N  Check: 3689  ..

Initial Score   =      5  Optimized Score  =      5  Significance  = -1.58
Residue Identity =    62%  Matches          =      5  Mismatches    =      3
Gaps            =      0  Conservative Substitutions               =      0

                                                        X      X
                                                        AACGTTCG
                                                         |||||

    GAAAGTAAAGATGCCTCCACGTCTGCAACATCTATAGACCAGTTGTGCAAGACGTTTAATCTTTCTTTGCAC
       110       120       130       140       150    X  160 X      170

    ACTCTGCAAATTCAGTGCGTGTTTTGCAGGAATGCA
     180       190       200       210
```

Mon Dec 3 08:02:28 2001

frag1-inv.res

```
> O <
O| |O  IntelliGenetics
> O <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file frag1-inv.res made by sdavid on Wed 28 Nov 101 14:18:57-PST.

Query sequence being compared:          FRAG1' (1-8)
                                                    6
Number of sequences searched:                       6
Number of scores above cutoff:

      Results of the initial comparison of FRAG1' (1-8) with:
        File : hpvcomplete.seq

   100-
     -
N    -
U  50-
M    -
B    -
E    -
R    -
     -
O    -
F  10-
     -
S    -                                                              *
E   5-
Q    -
U    -
E    -
N    -                                                    *
C    -
E    -
S    0-------------------------------------------------------------------
          |     |   |  |     ||     |    | |    ||     |     |     |
SCORE     0     1   |  1     2|     3    | 3    4|     5     5     6
STDEV            -7           -5          -2          0

                        PARAMETERS

Similarity matrix       Unitary          K-tuple                    4
Mismatch penalty              1          Joining penalty           30
Gap penalty                1.00          Window size                8
Gap size penalty           0.33
Cutoff score                  0
Randomization group           0

                     SEARCH STATISTICS

                        Mean        Median      Standard Deviation
Scores:                    5             7                    0.41

                         CPU                     Total Elapsed
Times:            00:00:00.00                  00:00:00.00

Number of residues:                     47269
Number of sequences searched:               6
Number of scores above cutoff:              6

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:
```

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
	**** 2 standard deviations above mean ****					
1. hpu06714	TOIG of: hpu06714 check: 486	7801	6	6	2.45	0
2. af131950	TOIG of: af131950 check: 557	7812	6	7	2.45	0
3. pph16	TOIG of: pph16 check: 6074	7904	6	7	2.45	0
4. al2360	TOIG of: al2360 check: 1580	7909	6	6	2.45	0
5. pph31a	TOIG of: pph31a check: 5866	7912	6	6	2.45	0
	**** 0 standard deviation from mean ****					
6. pph11	TOIG of: pph11 check: 3689	7931	5	6	0.00	0

1. FRAG1' (1-8)
   hpu06714 TOIG of: hpu06714 check: 4862 from: 1 to: 7801

TOIG of: hpu06714 check: 4862 from: 1 to: 7801

```
LOCUS       HPU06714      7801 bp    DNA                VRL       04-FEB-1997
DEFINITION  Human papillomavirus HPV-1A (3-3), complete genome.
ACCESSION   U06714
VERSION     U06714.1  GI:458704
KEYWORDS    .
SOURCE      Human papillomavirus.
  ORGANISM  Human papillomavirus
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 7801)
  AUTHORS   Danos,O., Katinka,M. and Yaniv,M.
  TITLE     Human papillomavirus 1A complete DNA sequence: a novel type of
            genome organization among Papovaviridae
  JOURNAL   EMBO J. 1, 231-236 (1982)
  MEDLINE   84182467
REFERENCE   2  (bases 1 to 7801)
  AUTHORS   Meissner,J.
  TITLE     Complete nucleotide sequencing of an HPV-1a variant and
            determination of extant errors in the prototype HPV-1a sequence
  JOURNAL   Virus Genes 9 (2), 189-191 (1995)
  MEDLINE   95250312
REFERENCE   3  (bases 1 to 7801)
  AUTHORS   Meissner,J.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-FEB-1994) John D. Meissner, Duke University Medical
            Center, Microbiology, 277 Carl Building, Research Drive, Durham, NC
            27710 USA
FEATURES             Location/Qualifiers
     source          1. .7801
                     /organism="Human papillomavirus"
                     /strain="HPV-1A (3-3)"
                     /db_xref="taxon:10566"
     variation       142
                     /replace="a"
     conflict        1283
                     /citation=[1]
                     /replace="a"
     conflict        2301
                     /citation=[1]
                     /replace="t"
     variation       2825
                     /replace="t"
     conflict        3884. .3886
                     /citation=[1]
                     /replace="aa"
     conflict        4332
                     /citation=[1]
                     /replace="a"
     conflict        4376. .4382
                     /citation=[1]
                     /replace="gagggaa"
     variation       5794
                     /replace="a"
     variation       6905
                     /replace="g"
```

```
     variation       7186. .7187
                     /note="15 bp deletion"
                     /citation=[1]
     variation       /replace="ctagattgtattgctat"
                     7560
     variation       /replace="c"
                     7618
     conflict        /replace="g"
                     7677. .7678
                     /citation=[1]
     variation       /replace="cc"
                     7787
                     /replace="g"
BASE COUNT      2389 a    1482 c    1664 g    2266 t
ORIGIN

HPU06714  Length: 7801  November 28, 2001 14:10  Type: N  Check: 4862  ..

Initial Score  =      6  Optimized Score  =      6  Significance  =  2.45
Residue Identity =  75%  Matches          =      6  Mismatches    =     2
Gaps           =      0  Conservative Substitutions               =     0

                                                      X       X
                                                      CGAACGTT
                                                      ||||||
   TCCACTGATGAATATGTAACCAGAACCAATCTCTTCTACCATGCAACATCTGAACGTCTACTGCTGGTCGGA
 5500       5510       5520       5530       5540       5550     X 5560

   CATCCTTTGTTTGAGATCTCCAGTAATCAAACTGTA
 5570       5580       5590       5600


2. FRAG1' (1-8)
   af131950      TOIG of: af131950  check: 557  from: 1  to: 7812

  TOIG of: af131950  check: 557  from: 1  to: 7812

LOCUS       AF131950      7812 bp     DNA              VRL       07-FEB-2001
DEFINITION  Human papillomavirus candHPV85, complete genome.
ACCESSION   AF131950
VERSION     AF131950.1  GI:4574720
KEYWORDS    .
SOURCE      Human papillomavirus candHPV85.
  ORGANISM  Human papillomavirus candHPV85
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 7812)
  AUTHORS   Chow,V.T.K. and Leong,P.W.F.
  TITLE     Complete nucleotide sequence, genomic organization and phylogenetic
            analysis of a novel genital human papillomavirus type, HLT7474-S
  JOURNAL   J. Gen. Virol. 80 (Pt 11), 2923-2929 (1999)
  MEDLINE   20047972
   PUBMED   10580054
REFERENCE   2  (bases 1 to 7812)
  AUTHORS   Chow,V.T.K. and Leong,W.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-FEB-1999) Department of Microbiology, National
            University of Singapore, 10 Kent Ridge Crescent, Singapore 119260,
            Singapore
FEATURES             Location/Qualifiers
     source          1. .7812
                     /organism="Human papillomavirus candHPV85"
                     /db_xref="taxon:151757"
                     /note="isolated from scraped uterine cervical cells from a
                     female sex worker; overlapping PCR products"
     protein_bind    40. .51
                     /note="putative"
                     /bound_moiety="E2 protein"
                     /function="transcriptional regulation"
     protein_bind    56. .67
                     /note="putative"
                     /bound_moiety="E2 protein"
                     /function="transcriptional regulation"
     TATA_signal     71. .79
     gene            /note="putative"
                     105. .578
     CDS             /gene="E6"
                     105. .578
                     /gene="E6"
                     /codon_start=1
                     /product="putative transforming protein E6"
                     /protein_id="AAD24181.1"
                     /db_xref="GI:4574721"
                     /translation="MAEFGNPATRPYKLPDLCNTLDTSLQDIEISCVYCKSVLQRTEV
                     YEFAFADLFVVYRDGIPYAACQNCLMFYSKIRELRYYSDSVYGETLEKLTNSNIYDIL
     gene            IRCLRCQKPLCPAEKLKHLNEKRRFHKIAGKYRGQCRRCMTRAQEQQGSRRETQV"
                     587. .913
     CDS             /gene="E7"
                     587. .913
                     /gene="E7"
                     /codon_start=1
                     /product="putative transforming protein E7"
                     /protein_id="AAD24182.1"
                     /db_xref="GI:4574722"
                     /translation="MHGPKPTVHDIVLDLEPYNEVQEVDLYCYEELNNSEEEIDEPDN
                     AINHRQPLLARREELQRHTICCVCCKCEASLQLVVESSAADLRDLQQLFLGTLSFLCP
     gene            LCAVLR"
                     920. .2866
     CDS             /gene="E1"
                     920. .2866
                     /gene="E1"
                     /codon_start=1
                     /product="putative replication protein E1"
                     /protein_id="AAD24183.1"
                     /db_xref="GI:4574723"
                     /translation="MADPEGTDGDGTGCNGWFFVQAIVDKKTGDTISEDEDEDGTDTG
                     SDLVDFIDTSNTNYMQADREAAQALLHAQEVETDTKLLHALKRKYGAHSTENSPCRDT
                     ASIHSNLSSPLQEISLNSYNNTAKRRLCSVPDSGYGNTQVETLQTQVTLDTNVFGDGK
                     NGDGLNSEACSTDNEMDIENQNPNSPMTQIVSLLKVNNKKAAILAKFKETYGLSFTDL
                     VRTFKSDKTTCTDWVAAICGVNPNIAEGFKTLIQPYVLYAHIQCMDCSWGVFILALLR
                     YKCGKNRLTVAKGLSTLLHVPDTHMLIEPPKLRSSCAALYWYRTGISNISEVTGDTPE
                     WIQRQTIIQHGIDDSVFDLSEMIQWAFDNDYIDESDIAYEYAQLADCNSNAAAFLKSN
                     CQAKYLRDCAVMCRHYKRAQRKQMNMSQWISYRCDKIDDGGDWKPIVQFLRFQGIEFI
                     TFLRAFKDFLKGTPKKNCIVIYGPANTGKSYFCMSLIQFLHGTVLSFVNSNSHFWLEP
                     LTDTKIAMVDDATPTCWSYFDNYMRNALDGNPISIDRKHKHLIQMKCPPMLITSNTNP
                     ATDDRWPYLRSRVTVFTFPHTFPFDSNGNPVYDINDKNWKCFFKRTWSRLDLHQEEED
     gene            TENDGCPIPTFKCVTGKDSRTLRD"
                     2799. .3932
     CDS             /gene="E2"
                     2799. .3932
                     /gene="E2"
                     /codon_start=1
                     /product="putative regulatory protein E2"
                     /protein_id="AAD24184.1"
                     /db_xref="GI:4574724"
                     /translation="MMDALSQRLNVLQEKILEHYETDSTDICDQIDYWKCVRLENAIY
                     YAAREHGLKTINHHVVPTFQISKSKAHEAIELQMALESLAKSEFKNELWTLQDTCQEL
                     YQTPPQQCFKKQGQTVEVRYDGDKDNTMHYTSWDYIYYVTEGDKWCKTKGYVNYCGLY
                     YIKEGQQTYYVQFKCDAQQYGQSGKWEVWYNGKKIECVNTCDKIIQCSESMYSTCDET
                     VSATAIARELQHPTTPYTEATTVCTQKSGGSAPTRNPFRHCGFTETSEVDGLSVDHLN
                     NPLLSAPAGNNGFRKHNSGDTTPIVHLKGDKNRLKCLRYRLQKFHNLYTQISCTWHWI
     gene            GGKETTRTGILTLAYANETQRQKFLDVVKIPNTVNVSLGYMTL"
                     <3412. .3699
     CDS             /gene="E4"
                     <3412. .3699
                     /gene="E4"
                     /note="start codon absent"
                     /codon_start=1
                     /product="putative protein E4"
                     /protein_id="AAD24185.1"
                     /db_xref="GI:4574725"
                     /translation="YNVLNLCTVPVTRQYPLLQLLENYNTPPHRIPKPPPCAPKRAGV
     gene            RRRLETHSDIVDSQKQVKSTDCPWTTSTTPCLVHLQATTGSGSTIVVTLRL"
                     3940. .4161
     CDS             /gene="E5"
                     3940. .4161
```

```
                     /gene="E5"
                     /codon_start=1
                     /product="putative protein E5"
                     /protein_id="AAD24186.1"
                     /db_xref="GI:4574726"
                     /translation="MLIAVFLLCFCVCLYVCCTIPLFPATNVCAHAWLLVFLLIVFIT
                     TPFTAFIVYIFFFILPMFLLHLHSVHTFD"
     gene            4226. .5629
                     /gene="L2"
     CDS             4226. .5629
                     /gene="L2"
                     /codon_start=1
                     /product="putative minor capsid protein L2"
                     /protein_id="AAD24187.1"
                     /db_xref="GI:4574727"
                     /translation="MVSHRAARRKRASATDLYKTCKQSGTCPPDVINKVEGTTLADKL
                     LQWTSLGIFLGGLGIGTGSGTGGRTGYIPLGGRPNTVVDVSPARPPVVIESVGPSDPS
                     IVTLVEESSIVTSGAPVPTFTGTSGFEITSSATTTPAVLDITPASGSVQLSSTSFTNP
                     AFTDPSVIEVPQTGEVSGDIFITTPTSGTHGYEEIPMHTFATQGRGTEPISSTPIPGV
                     RRVAGPRLYSQAYQQVKITNSDFISRPSTLVTFTNPAYEPIDTTLTFSPQDVVPDPDF
                     MDIVRLHRPALTSRRGTVRFSRLGKKLTMSTRSGKQIGAQVHYYHDISPISHIGESIE
                     MQPLLPDAAVTADTNGLFDIYADTDIDNNAMLYDRNISDVTQPTTSTISSVSSRYSNT
                     TIPLATSWDVPVHTGPDMTLPTTIPQWPNIVPLLPNNTHSVVLQGTNYYLWPNYYFIF
                     KKRKRVPYFLTDGFVAF"
     gene            5610. .7124
                     /gene="L1"
     CDS             5610. .7124
                     /gene="L1"
                     /codon_start=1
                     /product="putative major capsid protein L1"
                     /protein_id="AAD24188.1"
                     /db_xref="GI:4574728"
                     /translation="MALWRSSDSKVYLPPPSVAKVVNTDDYVTRTSTFYHAGSSRLLT
                     VGHPYYKVTSNGGRKQDIPKVSAYQYRVFRVTLPDPNKFGLPDTNIYNPETQRLVWAC
                     VGMEVGRGQPLGVSLSGHPFYNKLDDTENSHVATSVVTHDTRDNVSVDYKQTQLCIIG
                     CVPAIGEHWAKGTACKPGAVQTGDCPPLELVNTPIEDGDMIDTGYGAMDFSTLQDNKS
                     EVPLDICQSICKYPDYLQMSADAYGDSMFFCLRREQLFARHFWNRGGTIGDAVPETLY
                     IKGTNDRATPGSCIYSPSPSGSMVSSDAQMFNKPYWLHKAQGQNNGICWHNQLFITVV
                     DTTRSTNLTLSTATTNPVPSIYEPSKFKEYTRHVEEYDLQFIFQLCKITLTTDVMSYI
                     HNMDPTILDSWNFGVSPPPSASLVDTYRFLQSSAITCQKDVVVPQKKDPYEKLKFWNV
                     DLKEHFSSDLDQFPLGRKFLLQAGLRPKPTIGPRKRVASTSTATRPSKRKRTAK"
     protein_bind    7456. .7467
                     /note="putative"
                     /bound_moiety="E2 protein"
                     /function="transcriptional regulation"
     misc_feature    7647. .7666
                     /note="consensus motif of oncogenic anogenital HPV types"
     protein_bind    7776. .7787
                     /note="putative"
                     /bound_moiety="E2 protein"
                     /function="transcriptional regulation"
BASE COUNT     2502 a    1391 c    1558 g    2361 t
ORIGIN

AF131950  Length: 7812  November 28, 2001 14:10  Type: N  Check: 557  ..

Initial Score     =      6  Optimized Score  =      7  Significance =   2.45
Residue Identity  =    87%  Matches          =      7  Mismatches   =      1
Gaps              =      0  Conservative Substitutions               =      0

                                                  X       X
                                                  CGAACGTT
                                                  | ||||||
    TAGATTTGCACCAGGAAGAGGAGGATACAGAAAATGATGGATGCCCTATCCCAACGTTTAAATGTGTTACAG
     2770       2780       2790       2800       2810       2820  X    2830

    GAAAAGATTCTAGAACATTACGAGACTGATAGTACA
    2840       2850       2860       2870


3. FRAG1' (1-8)
   pph16        TOIG of: pph16  check: 6074  from: 1  to: 7904
```

```
TOIG of: pph16  check: 6074  from: 1  to: 7904

LOCUS       PPH16          7904 bp    DNA   circular  VRL      18-MAR-1994
DEFINITION  Human papillomavirus type 16 (HPV16), complete genome.
ACCESSION   K02718
VERSION     K02718.1  GI:333031
KEYWORDS    circular; complete genome.
SOURCE      Papilloma virus type 16 DNA, isolated from a human invasive
            cervical carcinoma.
  ORGANISM  Human papillomavirus type 16
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 7904)
  AUTHORS   Seedorf,K., Kraemmer,G., Duerst,M., Suhai,S. and Roewekamp,W.G.
  TITLE     Human papillomavirus type 16 DNA sequence
  JOURNAL   Virology 145, 181-185 (1985)
  MEDLINE   85246220
REFERENCE   2  (sites)
  AUTHORS   Kennedy,I.M., Haddow,J.K. and Clements,J.B.
  TITLE     A negative element in the human poapillomavirus type 16 genome acts
            at the level of late mRNA stability
  JOURNAL   J. Virol. 65, 2093-2097 (1991)
  MEDLINE   91162763
COMMENT     The sense strand of this double-stranded circular genome is shown,
            with a numbering system matching the first 60 bp of HPVa1, HPV6b
            and BPV1. The annotation of sites and features is solely based upon
            homology comparison with these other papillomaviruses. In addition
            to the coding sequences reported below, the authors note open
            reading frames which do not start with 'ATG', but which are found
            in other papillomaviruses. In particular, a second portion of the
            E1 gene may be located out to base 2813 (the E1 protein is thought
            to be generally involved in DNA replication).
            A potential 'CAT'-box region is found beginning at base 7895 below,
            and 'TATA' boxes for early and late transcripts may be located at
            17, 65 and 4289. Potential polyadenylation signals are at bases
            4213 and 7260.
            HPV16, in comparison to HPV types 6 and 11, is more often
            associated with malignant genital cancers in humans.
FEATURES             Location/Qualifiers
     source          1. .7904
                     /organism="Human papillomavirus type 16"
                     /db_xref="taxon:10581"
     TATA_signal     17. .23
     TATA_signal     65. .71
     gene            83. .559
                     /gene="E6"
     CDS             83. .559
                     /gene="E6"
                     /note="E6 ORF from 65 to 559; putative"
                     /codon_start=1
                     /product="transforming protein"
                     /protein_id="AAA46939.1"
                     /db_xref="GI:333032"
                     /translation="MHQKRTAMFQDPQERPRKLPQLCTELQTTIHDIILECVYCKQQL
                     LRREVYDFAFRDLCIVYRDGNPYAVCDKCLKFYSKISEYRHYCYSLYGTTLEQQYNKP
                     LCDLLIRCINCQKPLCPEEKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL"
     gene            562. .858
                     /gene="E7"
     CDS             562. .858
                     /gene="E7"
                     /note="E7 ORF from 544 to 858; putative"
                     /codon_start=1
                     /product="transforming protein"
                     /protein_id="AAA46940.1"
                     /db_xref="GI:333033"
                     /translation="MHGDTPTLHEYMLDLQPETTDLYCYEQLNDSSEEEDEIDGPAGQ
                     AEPDRAHYNIVTFCCKCDSTLRLCVQSTHVDIRTLEDLLMGTLGIVCPICSQKP"
     gene            join(865. .1140,1140. .2813)
                     /gene="E1"
     CDS             join(865. .1140,1140. .2813)
                     /gene="E1"
                     /note="E1 interrupted ORF from 859 to 2813; putative"
                     /codon_start=1
```

```
                     /product="replication protein"
                     /protein_id="AAA46936.1"
                     /db_xref="GI:459912"
                     /translation="MADPAGTNGEEGTGCNGWFYVEAVVEKKTGDAISDDENENDSDT
                     GEDLVDFIVNDNDYLTQAETETAHALFTAQEAKQHRDAVQVLKRKYLVSPLSDISGCV
                     DNNISPRLKAICIEKQSRAAKRRLFESEDSGYGNTEVETQQMLQVEGRHETETPCSQY
                     SGGSGGGCSQYSSGSGGEGVSERHTICQTPLTNILNVLKTSNAKAAMLAKFKELYGVS
                     FSELVRPFKSNKSTCCDWCIAAFGLTPSIADSIKTLLQQYCLYLHIQSLACSWGMVVL
                     LLVRYKCGKNRETIEKLLSKLLCVSPMCMMIEPPKLRSTAAALYWYKTGISNISEVYG
                     DTPEWIQRQTVLQHSFNDCTFELSQMVQWAYDNDIVDDSEIAYKYAQLADTNSNASAF
                     LKSNSQAKIVKDCATMCRHYKRAEKKQMSMSQWIKYRCDRVDDGGDWKQIVMFLRYQG
                     VEFMSFLTALKRFLQGIPKKNCILLYGAANTGKSLFGMSLMKFLQGSVICFVNSKSHF
                     WLQPLADAKIGMLDDATVPCWNYIDDNLRNALDGNLVSMDVKHRPLVQLKCPPLLITS
                     NINAGTDSRWPYLHNRLVVFTFPNEFPFDENGNPVYELNDKNWKSFFSRTWSRLSLHE
                     DEDKENDGDSLPTFKCVSGQNTNTL"
gene                 2755. .3852
CDS                  2755. .3852
                     /gene="E2"
                     /note="E2 ORF from 2725 to 3852; putative"
                     /codon_start=1
                     /product="regulatory protein"
                     /protein_id="AAA46941.1"
                     /db_xref="GI:333035"
                     /translation="METLCQRLNVCQDKILTHYENDSTDLRDHIDYWKHMRLECAIYY
                     KAREMGFKHINHQVVPTLAVSKNKALQAIELQLTLETIYNSQYSNEKWTLQDVSLEVY
                     LTAPTGCIKKHGYTVEVQFDGDICNTMHYTNWTHIYICEEASVTVVEGQVDYYGLYYV
                     HEGIRTYFVQFKDDAEKYSKNKVWEVHAGGQVILCPTSVFSSNEVSSPEIIRQHLANH
                     PAATHTKAVALGTEETQTTIQRPRSEPDTGNPCHTTKLLHRDSVDSAPILTAFNSSHK
                     GRINCNSNTTPIVHLKGDANTLKCLRYRFKKHCTLYTAVSSTWHWTGHNVKHKSAIVT
gene                 LTYDSEWQRDQFLSQVKIPKTITVSTGFMSI"
CDS                  3332. .3619
                     /gene="E4"
                     <3332. .3619
                     /gene="E4"
                     /note="E4 ORF from 3332 to 3619; putative"
                     /codon_start=1
                     /protein_id="AAA46937.1"
                     /db_xref="GI:459913"
                     /translation="YYVLHLCLAATKYPLLKLLGSTWPTTPPRPIPKPSPWAPKKHRR
gene                 LSSDQDQSQTPETPATPLSCCTETQWTVLQSSLHLTAHTKDGLTVIVTLHP"
CDS                  3863. .4099
                     /gene="E5"
                     <3863. .4099
                     /gene="E5"
                     /note="E5 ORF from 3863 to 4099; putative"
                     /codon_start=1
                     /protein_id="AAA46938.1"
                     /db_xref="GI:459914"
                     /translation="YCIHNITGVLFALLCVLLCVCLLIRPLLLSVSTYTSLIILVLLL
polyA_signal         WITAASAFRCFIVYIIFVYIPLFLIHTHARFLIT"
                     4213. .4218
gene                 /note="putative"
CDS                  4235. .5656
                     /gene="L2"
                     4235. .5656
                     /gene="L2"
                     /note="L2 ORF from 4133 to 5656; putative"
                     /codon_start=1
                     /product="minor capsid protein"
                     /protein_id="AAA46942.1"
                     /db_xref="GI:333036"
                     /translation="MRHKRSAKRTKRASATQLYKTCKQAGTCPPDIIPKVEGKTIAEQ
                     ILQYGSMGVFFGGLGIGTGSGTGGRTGYIPLGTRPPTATDTLAPVRPPLTVDPVGPSD
                     PSIVSLVEETSFIDAGAPTSVPSIPPDVSGFSITTSTDTTPAILDINNTVTTVTTHNN
                     PTFTDPSVLQPPTPAETGGHFTLSSSTISTHNYEEIPMDTFIVSTNPNTVTSSTPIPG
                     SRPVARLGLYSRTTQQVKVVDPAFVTTPTKLITYDNPAYEGIDVDNTLYFSSNDNSIN
                     IAPDPDFLDIVALHRPALTSRRTGIRYSRIGNKQTLRTRSGKSIGAKVHYYYDLSTID
                     PAEEIELQTITPSTYTTTSHAASPTSINNGLYDIYADDFITDTSTTPVPSVPSTSLSG
                     YIPANTTIPFGGAYNIPLVSGPDIPINITDQAPSLIPIVPGSPQYTIIADAGDFYLHP
TATA_signal          SYYMLRKRRKRLPYFFSDVSLAA"
                     4289. .4295
                     /gene="L2"
gene                 5559. .7154
                     /gene="L1"
CDS                  5559. .7154
                     /gene="L1"
                     /note="L1 ORF from 5526 to 7154; putative"
                     /codon_start=1
                     /product="major capsid protein"
                     /protein_id="AAA46943.1"
                     /db_xref="GI:333037"
                     /translation="MQVTFIYILVITCYENDVNVYHIFFQMSLWLPSEATVYLPPVPV
                     SKVVSTDEYVARTNIYYHAGTSRLLAVGHPYFPIKKPNNNKILVPKVSGLQYRVFRIH
                     LPDPNKFGFPDTSFYNPDTQRLVWACVGVEVGRGQPLGVGISGHPLLNKLDDTENASA
                     YAANAGVDNRECISMDYKQTQLCLIGCKPPIGEHWGKGSPCTNVAVNPGDCPPLELIN
                     TVIQDGDMVHTGFGAMDFTTLQANKSEVPLDICTSICKYPDYIKMVSEPYGDSLFFYL
                     RREQMFVRHLFNRAGTVGENVPDDLYIKGSGSTANLASSNYFPTPSGSMVTSDAQIFN
                     KPYWLQRAQGHNNGICWGNQLFVTVVDTTRSTNMSLCAAISTSETTYKNTNFKEYLRH
                     GEEYDLQFIFQLCKITLTADVMTYIHSMNSTILEDWNFGLQPPPGGTLEDTYRFVTQA
                     IACQKHTPPAPKEDDPLKKYTFWEVNLKEKFSADLDQFPLGRKFLLQAGLKAKPKFTL
                     GKRKATPTTSSTSTTAKRKKRKL"
polyA_signal         7260. .7265
BASE COUNT      2601 a    1377 c    1509 g    2417 t
ORIGIN      Unreported.

  PPH16  Length: 7904  November 28, 2001 14:10  Type: N  Check: 6074  ..

Initial Score   =      6  Optimized Score  =      7  Significance  =   2.45
Residue Identity =    87%  Matches         =      7  Mismatches    =      1
Gaps            =      0  Conservative Substitutions                =      0

                                                        X       X
                                                        CGAACGTT
                                                        | ||||||
    GATTAAGTTTGCACGAGGACGAGGACAAGGAAAACGATGGAGACTCTTTGCCAACGTTTAAATGTGTGTCAG
2720       2730       2740       2750       2760       2770       2780       2790

    GACAAAATACTAACACATTATGAAAATGATAGTACA
         2800       2810       2820


4. FRAG1' (1-8)
   a12360       TOIG of: a12360  check: 1580  from: 1  to: 7909

  TOIG of: a12360  check: 1580  from: 1  to: 7909

LOCUS       A12360          7909 bp    DNA                     PAT       12-DEC-1993
DEFINITION  complete nucleotide sequence of HPV-33.
ACCESSION   A12360
VERSION     A12360.1  GI:492936
KEYWORDS    .
SOURCE      Human papillomavirus type 33.
  ORGANISM  Human papillomavirus type 33
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 7909)
  AUTHORS   .
  JOURNAL   Patent: WO 8705630-A 1 24-SEP-1987;
FEATURES             Location/Qualifiers
     source          1. .7909
                     /organism="Human papillomavirus type 33"
                     /db_xref="taxon:10586"
BASE COUNT      2544 a    1354 c    1535 g    2474 t      2 others
ORIGIN

A12360  Length: 7909  November 28, 2001 14:10  Type: N  Check: 1580  ..

Initial Score   =      6  Optimized Score  =      6  Significance  =   2.45
Residue Identity =    75%  Matches         =      6  Mismatches    =      2
Gaps            =      0  Conservative Substitutions                =      0

                                                        X       X
                                                        CGAACGTT
                                                        ||||||
    GTTCACAAAGTGCTGCGGAGGACGTTGTTGATCGTGCTGCAAACCCGTGTAGAACGTCTATTAATAAAAATA
```

```
      1160       1170       1180       1190       1200  X    1210       1220

 AAGAATGCACATACAGAAAACGAAAAATAGATGAGC
   1230       1240       1250       1260


5. FRAG1' (1-8)
   pph31a        TOIG of: pph31a  check: 5866  from: 1  to: 7912

 TOIG of: pph31a  check: 5866  from: 1  to: 7912

LOCUS       PPH31A       7912 bp    DNA   circular  VRL       18-MAR-1994
DEFINITION  Human papillomavirus type 31 (HPV-31) complete genome.
ACCESSION   J04353
VERSION     J04353.1  GI:333048
KEYWORDS    complete genome.
SOURCE      Human papillomavirus type 31 DNA.
  ORGANISM  Human papillomavirus type 31
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 7912)
  AUTHORS   Goldsborough,M.D., DiSelvestre,D., Temple,G.F. and Lorincz,A.T.
  TITLE     Nucleotide sequence of human papillomavirus type 31: A cervical
            neoplasia associated virus
  JOURNAL   Virology 171, 306-311 (1989)
  MEDLINE   89299478
COMMENT     Draft entry and computer-readable copy of sequence [1] kindly
            submitted by M.D.Goldsborough, 05-JUL-1989.
FEATURES             Location/Qualifiers
     source          1. .7912
                     /organism="Human papillomavirus type 31"
                     /db_xref="taxon:10585"
     TATA_signal     19. .24
     TATA_signal     69. .74
     gene            108. .557
                     /gene="E6"
     CDS             108. .557
                     /gene="E6"
                     /note="ORF E6 from bp 39 to 557"
                     /codon_start=1
                     /product="transforming protein"
                     /protein_id="AAA46950.1"
                     /db_xref="GI:459916"
                     /translation="MFKNPAERPRKLHELSSALEIPYDELRLNCVYCKGQLTETEVLD
                     FAFTDLTIVYRDDTPHGVCTKCLRFYSKVSEFRWYRYSVYGTTLEKLTNKGICDLLIR
                     CITCQRPLCPEEKQRHLDKKKRFHNIGGRWTGRCIACWRRPRTETQV"
     misc_feature    228. .236
                     /gene="E6"
                     /standard_name="Splice donor"
     misc_feature    403. .414
                     /gene="E6"
                     /standard_name="Splice acceptor"
     gene            560. .856
                     /gene="E7"
     CDS             560. .856
                     /gene="E7"
                     /note="ORF E7 from bp 545 to 856"
                     /codon_start=1
                     /product="transforming protein"
                     /protein_id="AAA46951.1"
                     /db_xref="GI:459917"
                     /translation="MRGETPTLQDYVLDLQPEATDLHCYEQLPDSSDEEDVIDSPAGQ
                     AEPDTSNYNIVTFCCQCKSTLRLCVQSTQVDIRILQELLMGSFGIVCPNCSTRL"
     gene            862. .2751
                     /gene="E1"
     CDS             862. .2751
                     /gene="E1"
                     /note="ORF E1 from bp 850 to bp 2751"
                     /codon_start=1
                     /product="replication protein"
                     /protein_id="AAA46952.1"
                     /db_xref="GI:459918"
                     /translation="MADPAGTDGEGTGCNGWFYVEAVIDRQTGDNISEDENEDSSDTG
                     EDMVDFIDNCNVYNNQAEAETAQALFHAQEAEEHAEAVQVLKRKYVGSPLSDISSCVD
                     YNISPRLKAICIENNSKTAKRRLFELPDSGYGNTEVETQQMVQVEEQQTTLSCNGSDG
                     THSERENETPTRNILQVLKTSNGKAAMLGKFKELYGVSFMELIRPFQSNKSTCTDWCV
                     AAFGVTGTVAEGFKTLLQPYCLYCHLQSLACSWGMVMLMLVRFKCAKNRITIEKLLEK
                     LLCISTNCMLIQPPKLRSTAAALYWYRTGMSNISDVYGETPEWIERQTVLQHSFNDTT
                     FDLSQMVQWAYDNDVMDDSEIAYKYAQLADSDSNACAFLKSNSQAKIVKDCGTMCRHY
                     KRAEKRQMSMGQWIKSRCDKVSDEGDWRDIVKFLRYQQIEFVSFLSALKLFLKGVPKK
                     NCILIHGAPNTGKSYFGMSLISFLQGCIISYANSKSHFWLQPLADAKIGMLDDATTPC
                     WHYIDNYLRNALDGNPVSIDVKHKALMQLKCPPLLITSNINAGKDDRWPYLHSRLVVF
                     TFPNPFPFDKNGNPVYELSDKNWKSFFSRTWCRLNLHEEEDKENDGDSFSTFKCVSGQ
                     NIRTL"
     gene            2693. .3811
                     /gene="E2"
     CDS             2693. .3811
                     /gene="E2"
                     /note="ORF E2 from bp 2663 to 3811"
                     /codon_start=1
                     /product="regulatory protein"
                     /protein_id="AAA46953.1"
                     /db_xref="GI:459919"
                     /translation="METLSQRLNVCQDKILEHYENDSKRLCDHIDYWKHIRLECVLMY
                     KAREMGIHSINHQVVPALSVSKAKALQAIELQMMLETLNNTEYKNEDWTMQQTSLELY
                     LTAPTGCLKKHGYTVEVQFDGDVHNTMHYTNWKFIYLCIDGQCTVVEGQVNCKGIYYV
                     HEGHITYFVNFTEEAKKYGTGKKWEVHAGGQVIVFPESVFSSDEISFAGIVTKLPTAN
                     NTTTSNSKTCALGTSEGVRRATTSTKRPRTEPEHRNTHHPNKLLRGDSVDSVNCGVIS
                     AAACTNQTRAVSCPATTPIIHLKGDANILKCLRYRLSKYKQLYEQVSSTWHWTCTDGK
                     HKNAIVTLTYISTSQRDDFLNTVKIPNTVSVSTGYMTI"
     gene            3270. .3578
                     /gene="E4"
     CDS             <3270. .3578
                     /gene="E4"
                     /note="ORF E4 from bp 3270 to 3578"
                     /codon_start=1
                     /protein_id="AAA46949.1"
                     /db_xref="GI:459915"
                     /translation="LFFLNLYLAVTKYPLLGLLQSYQQPTTPPHRIPKPAPWAPVKVC
                     GGRRRLLSDQEQSQSTETPTTPTSCCEATPWTVSTVGLSVQLHAQTKQGLSVVLQLHL
                     "
     gene            3816. .4070
                     /gene="E5"
     CDS             3816. .4070
                     /gene="E5"
                     /note="ORF E5 from bp 3804 to 4070"
                     /codon_start=1
                     /protein_id="AAA46954.1"
                     /db_xref="GI:459920"
                     /translation="MIELNISTVSIVLCFLLCFCVLLFVCLVIRPLVLSVSVYATLLL
                     LIVILWVIATSPLRCFCIYVVFIYIPLFVIHTHASFLSQQ"
     repeat_region   4099. .4134
     polyA_signal    4138. .4143
                     /note="putative"
     repeat_region   4143. .4158
     gene            4171. .5571
                     /gene="L2"
     CDS             4171. .5571
                     /gene="L2"
                     /note="ORF L2 from bp 4060 to bp 5571"
                     /codon_start=1
                     /product="minor capsid protein"
                     /protein_id="AAA46955.1"
                     /db_xref="GI:459921"
                     /translation="MRSKRSTKRTKRASATQLYQTCKAAGTCPSDVIPKIEHTTIADQ
                     ILRYGSMGVFFGGLGIGSGSGTGGRTGYVPLSTRPSTVSEASIPIRPPVSIDPVGPLD
                     PSIVSLVEESGIVDVGAPAPIPHPPTTSGFDIATTADTTPAILDVTSVSTHENPTFTD
                     PSVLQPPTPAETSGHLLLSSSSISTHNYEEIPMDTFIVSTNNENITSSTPIPGVRRPA
                     RLGLYSKATQQVKVIDPTFLSAPKQLITYENPAYETVNAEESLYFSNTSHNIAPDPDF
                     LDIIALHRPALTSRRNTVRYSRLGNKQTLRTRSGATIGARVHYYYDISSINPAGESIE
                     MQPLGASATTTSTLNDGLYDIYADTDFTVDTPATHNVSPSTAVQSTSAVSAYVPTNTT
                     VPLSTGFDIPIFSGPDVPIEHAPTQVFPFPLAPTTPQVSIFVDGGDFYLHPSYYMLKR
                     RRKRVSYFFTDVSVAA"
     gene            5552. .7066
                     /gene="L1"
     CDS             5552. .7066
```

```
                     /gene="L1"
                     /note="ORF L1 from bp 5516 to  7066"
                     /codon_start=1
                     /product="major capsid protein"
                     /protein_id="AAA46956.1"
                     /db_xref="GI:459922"
                     /translation="MSLWRPSEATVYLPPVPVSKVVSTDEYVTRTNIYYHAGSARLLT
                     VGHPYYSIPKSDNPKKIVVPKVSGLQYRVFRVRLPDPNKFGFPDTSFYNPETQRLVWA
                     CVGLEVGRGQPLGVGISGHPLLNKFDDTENSNRYAGGPGTDNRECISMDYKQTQLCLL
                     GCKPPIGEHWGKGSPCSNNAITPGDCPPLELKNSVIQDGDMVDTGFGAMDFTALQDTK
                     SNVPLDICNSICKYPDYLKMVAEPYGDTLFFYLRREQMFVRHFFNRSGTVGESVPTDL
                     YIKGSGSTATLANSTYFPTPSGSMVTSDAQIFNKPYWMQRAQGHNNGICWGNQLFVTV
                     VDTTRSTNMSVCAAIANSDTTFKSSNFKEYLRHGEEFDLQFIFQLCKITLSADIMTYI
                     HSMNPAILEDWNFGLTTPPSGSLEDTYRFVTSQAITCQKTAPQKPKEDPFKDYVFWEV
                     NLKEKFSADLDQFPLGRKFLLQAGYRARPKFKAGKRSAPSASTTTPAKRKKTKK"
     polyA_signal    7227. .7231
     repeat_region   7291. .7302
     repeat_region   7314. .7333
                     /note="putative"
     protein_bind    7406. .7420
                     /standard_name="glucocorticoid responsive element"
     protein_bind    /bound_moiety="hormone receptor"
                     7477. .7488
                     /function="gene transcription"
     enhancer        /bound_moiety="E2"
                     7542. .7549
     protein_bind    /standard_name="keratinocyte-dependant enhancer"
                     7868. .7879
                     /function="gene transcription"
                     /bound_moiety="E2"
BASE COUNT      2528 a    1364 c    1572 g    2448 t
ORIGIN

PPH31A  Length: 7912  November 28, 2001 14:10  Type: N  Check: 5866  ..

Initial Score     =       6  Optimized Score  =       6  Significance  =  2.45
Residue Identity  =     75%  Matches          =       6  Mismatches    =       2
Gaps              =       0  Conservative Substitutions                =       0

                                                      X       X
                                                      CGAACGTT
                                                      ||||||
    GATTAAATTTGCACGAGGAAGAGGACAAAGAAAACGATGGAGACTCTTTCTCAACGTTTAAATGTGTGTCAG
    2660       2670       2680       2690       2700       2710    X  2720

    GACAAAATATTAGAACATTATGAAAATGATAGTAAA
2730       2740       2750       2760

6. FRAG1' (1-8)
   pph11          TOIG of: pph11  check: 3689  from: 1  to: 7931

TOIG of: pph11  check: 3689  from: 1  to: 7931

LOCUS       PPH11          7931 bp    DNA    circular  VRL       02-JUN-1994
DEFINITION  Human papillomavirus type 11 (HPV-11) complete genome.
ACCESSION   M14119
VERSION     M14119.1  GI:333026
KEYWORDS    complete genome.
SOURCE      Human laryngeal papillomavirus type 11 DNA.
  ORGANISM  Human papillomavirus type 11
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 7931)
  AUTHORS   Dartmann,K., Schwarz,E., Gissmann,L. and zur Hausen,H.
  TITLE     The nucleotide sequence and genome organization of human papilloma
            virus type 11
  JOURNAL   Virology 151, 124-130 (1986)
  MEDLINE   86181601
COMMENT     ORF L1 is assumed to encode the major structural protein.
FEATURES             Location/Qualifiers
     source          1. .7931
                     /organism="Human papillomavirus type 11"
                     /db_xref="taxon:10580"
     CAAT_signal     9. .15
                     /note="putative"
     protein_bind    35. .46
                     /note="putative"
                     /function="gene transcription"
     protein_bind    /bound_moiety="E2"
                     50. .61
                     /note="putative"
                     /function="gene transcription"
     TATA_signal     /bound_moiety="E2"
                     66. .71
                     /note="putative"
     gene            102. .554
     CDS             /gene="E6"
                     102. .554
                     /gene="E6"
                     /note="102 is position of first start codon in ORF E6;
                     putative"
                     /codon_start=1
                     /product="transforming protein"
                     /protein_id="AAA46927.1"
                     /db_xref="GI:496193"
                     /translation="MESKDASTSATSIDQLCKTFNLSLHTLQIQCVFCRNALTTAEIY
                     AYAYKNLKVVWRDNFPFAACACCLELQGKINQYRHFNYAAYAPTVEEETNEDILKVLI
     gene            RCYLCHKPLCEIEKLKHILGKARFIKLNNQWKGRCLHCWTTCMEDLLP"
                     530. .826
     CDS             /gene="E7"
                     530. .826
                     /gene="E7"
                     /note="530 is position of first start codon in ORF E7;
                     putative"
                     /codon_start=1
                     /product="transforming protein"
                     /protein_id="AAA46928.1"
                     /db_xref="GI:496194"
                     /translation="MHGRLVTLKDIVLDLQPPDPVGLHCYEQLEDSSEDEVDKVDKQD
     gene            AQPLTQHYQILTCCCGCDSNVRLVVECTDGDIRQLQDLLLGTLNIVCPICAPKP"
                     832. .2781
     CDS             /gene="E1"
                     832. .2781
                     /gene="E1"
                     /note="832 is position of first start codon in ORF E1;
                     putative"
                     /codon_start=1
                     /product="replication protein"
                     /protein_id="AAA46929.1"
                     /db_xref="GI:496195"
                     /translation="MADDSGTENEGSGCTGWFMVEAIVEHTTGTQISEDEEEEVEDSG
                     YDMVDFIDDRHITQNSVEAQALFNRQEADAHYATVQDLKRKYLGSPYVSPISNVANAV
                     ESEISPRLDAIKLTTQPKKVKRRLFETRELTDSGYGYSEVEAATQVEKHGDPENGGDG
                     QERDTGRDIEGEGVEHREAEAVDDSTREHADTSGILELLKCKDIRSTLHGKFKDCFGL
                     SFVDLIRPFKSDRTTCADWVVAGFGIHHSIADAFQKLIEPLSLYAHIQWLTNAWGMVL
                     LVLIRFKVNKSRCTVARTLGTLLNIPENHMLIEPPKIQSGVRALYWFRTGISNASTVI
                     GEAPEWITRQTVIEHSLADSQFKLTEMVQWAYDNDICEESEIAFEYAQRGDFDSNARA
                     FLNSNMQAKYVKDCAIMCRHYKHAEMKKMSIKQWIKYRGTKVDSVGNWKPIVQFLRHQ
                     NIEFIPFLSKLKLWLHGTPKKNCIAIVGPPDTGKSCFCMSLIKFLGGTVISYVNSCSH
                     FWLQPLTDAKVALLDDATQPCWTYMDTYMRNLLDGNPMSIDRKHRALTLIKCPPLLVT
                     SNIDISKEEKYKYLHSRVTTFTFPNPFPFDRNGNAVYELSDANWKCFFERLSSSLDIE
     gene            DSEDEEDGSNSQAFRCVPGSVVRTL"
                     2723. .3826
     CDS             /gene="E2"
                     2723. .3826
                     /gene="E2"
                     /note="2723 is position of first start codon in ORF E2;
                     putative"
                     /codon_start=1
                     /product="regulatory protein"
                     /protein_id="AAA46930.1"
                     /db_xref="GI:496196"
                     /translation="MEAIAKRLDACQDQLLELYEENSIDIHKHIMHWKCIRLESVLLH
                     KAKQMGLSHIGLQVVPPLTVSETKGHNAIEMQMHLESLAKTQYGVEPWTLQDTSYEMW
```

```
                LTPPKRCFKKQGNTVEVKFDGCEDNVMEYVVWTHIYLQDNDSWVKVTSSVDAKGIYYT
                CGQFKTYYVNFNKEAQKYGSTNHWEVCYGSTVICSPASVSSTVREVSIAEPTTYTPAQ
                TTAPTVSACTTEDGVSAPPRKRARGPSTNNTLCVANIRSVDSTINNIVTDNYNKHQRR
                NNCHSAATPIVQLQGDSNCLKCFRYRLNDKYKHLFELASSTWHWASPEAPHKNAIVTL
                TYSSEEQRQQFLNSVKIPPTIRHKVGFMSLHLL"
gene            3255. .3581
                /gene="E4"
CDS             3255. .3581
                /gene="E4"
                /note="3255 is position of first start codon in ORF E4;
                putative"
                /codon_start=1
                /protein_id="AAA46931.1"
                /db_xref="GI:496197"
                /translation="MVVPIIGKYVMAAQLYVLLHLYLALYEKYPLLNLLHTPPHRPPP
                LQCPPAPRKTACRRRLGSEHVDRPLTTPCVWPTSDPWTVQSTTSSLTITTSTKEGTTV
                TVQLRL"
gene            3871. .4146
                /gene="E5A"
CDS             3871. .4146
                /gene="E5A"
                /note="3871 is position of first start codon in ORFE5A;
                putative"
                /codon_start=1
                /protein_id="AAA46932.1"
                /db_xref="GI:496198"
                /translation="MEVVPVQIAAATTTTLILPVVIAFAVCILSIVLIILISDFVVYT
                SVLVLTLLLYLLLWLLLTTPLQFFLLTLCVCYFPAFYIHIYIVQTQQ"
gene            4146. .4370
                /gene="E5B"
CDS             4146. .4370
                /gene="E5B"
                /note="4146 is position of first start codon in ORF E5B;
                putative"
                /codon_start=1
                /protein_id="AAA46933.1"
                /db_xref="GI:496199"
                /translation="MVMLTCHLNDGDTWLFLWLFTAFVVAVLGLLLLHYRAVHGTEKT
                KCAKCKSNRNTTVDYVYMSHGDNGDYVYMN"
polyA_signal    4371. .4376
                /note="putative"
gene            4417. .5784
                /gene="L2"
CDS             4417. .5784
                /gene="L2"
                /note="4417 is position of first start codon in ORF L2;
                putative"
                /codon_start=1
                /product="minor capsid protein"
                /protein_id="AAA46934.1"
                /db_xref="GI:496200"
                /translation="MKPRARRRKRASATQLYQTCKATGTCPPDVIPKVEHTTIADQIL
                KWGSLGVFFGGLGIGTGAGSGGRAGYIPLGSSPKPAITGGPAARPPVLVEPVAPSDPS
                IVSLIEESAIINAGAPEVVPPTQGGFTITSSESTTPAILDVSVTNHTTTSVFQNPLFT
                EPSVIQPQPPVEASGHILISAPTITSQHVEDIPLDTFVVSSSDSGPTSSTPLPRAFPR
                PRVGLYSRALQQVQVTDPAFLSTPQRLVTYDNPVYEGEDVSLQFTHESIHNAPDEAFM
                DIIRLHRPAITSRRGLVRFSRIGQRGSMYTRSGQHIGARIHYFQDISPVTQAAEEIEL
                HPLVAAENDTFDIYAEPFDPIPDPVQHSVTQSYLTSTPNTLSQSWGNTTVPLSIPSDW
                FVQSGPDITFPTASMGTPFSPVTPALPTGPVFITGSDFYLHPTWYFARRRRKRIPLFF
                TDVAA"
polyA_signal    4545. .4551
                /gene="L2"
                /note="putative"
gene            5771. .7276
                /gene="L1"
CDS             5771. .7276
                /gene="L1"
                /note="5771 is position of first start codon in ORF L1;
                putative"
                /codon_start=1
                /product="major capsid protein"
                /protein_id="AAA46935.1"
                /db_xref="GI:496201"
                /translation="MWRPSDSTVYVPPPNPVSKVVATDAYVKRTNIFYHASSSRLLAV
                GHPYYSIKKVNKTVVPKVSGYQYRVFKVVLPDPNKFALPDSSLFDPTTQRLVWACTGL
                EVGRGQPLGVGVSGHPLLNKYDDVENSGGYGGNPGQDNRVNVGMDYKQTQLCMVGCAP
                PLGEHWGKGTQCSNTSVQNGDCPPLELITSVIQDGDMVDTGFGAMNFADLQTNKSDVP
                LDICGTVCKYPDYLQMAADPYGDRLFFYLRKEQMFARHFFNRAGTVGEPVPDDLLVKG
                GNNRSSVASSIYVHTPSGSLVSSEAQLFNKPYWLQKAQGHNNGICWGNHLFVTVVDTT
                RSTNMTLCASVSKSATYTNSDYKEYMRHVEEFDLQFIFQLCSITLSAEVMAYIHTMNP
                SVLEDWNFGLSPPPNGTLEDTYRYVQSQAITCQKPTPEKEKQDPYKDMSFWEVNLKEK
                FSSELDQFPLGRKFLLQSGYRGRTSARTGIKRPAVSKPSTAPKRKRTKTKK"
5'UTR           join(7277. .7931,1. .101)
                /standard_name="LCR"
                /function="regulatory region"
repeat_region   7320. .7360
                /standard_name="direct repeats"
                /note="putative"
repeat_region   7339. .7374
                /standard_name="direct repeats"
                /note="putative"
repeat_region   7374. .7403
                /note="putative"
polyA_signal    7457. .7462
                /note="putative"
protein_bind    7592. .7603
                /note="putative"
                /function="gene transcription"
                /bound_moiety="E2"
polyA_signal    7748. .7753
                /note="putative"
protein_bind    7890. .7901
                /note="putative"
                /function="gene transcription"
                /bound_moiety="E2"
BASE COUNT      2406 a   1519 c   1736 g   2270 t
ORIGIN      4557 bp upstream of HindIII site.

  PPH11  Length: 7931  November 28, 2001 14:10  Type: N  Check: 3689  ..

Initial Score   =      5  Optimized Score  =      6  Significance  =  0.00
Residue Identity =    77%  Matches          =      7  Mismatches    =     1
Gaps            =      1  Conservative Substitutions               =     0

                                                          X       X
                                                          CGA-ACGTT
                                                          | | |||||

ATGGAAAGTAAAGATGCCTCCACGTCTGCAACATCTATAGACCAGTTGTGCAAGACGTTTAATCTTTCTTTG
      110       120       130       140       150 X     160       170

CACACTCTGCAAATTCAGTGCGTGTTTTGCAGGAATG
   180       190       200       210
```

Mon Dec 3 08:02:34 2001

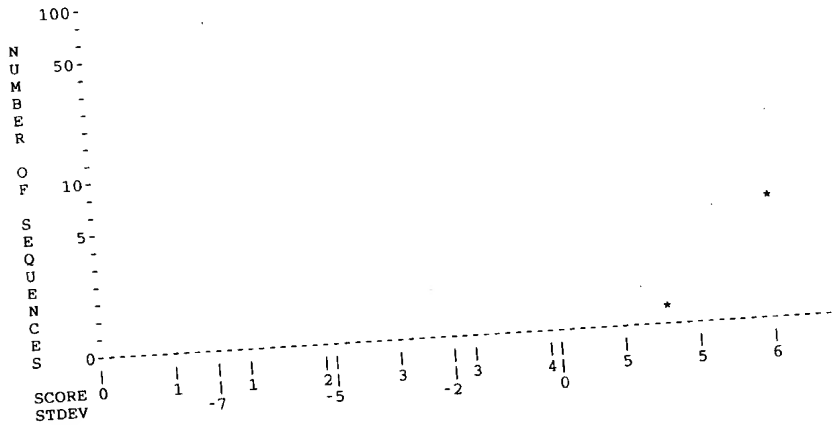
> O <
O| |O IntelliGenetics
> O <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file frag2.res made by sdavid on Wed 28 Nov 101 14:19:54-PST.

Query sequence being compared: FRAG2 (1-8)
Number of sequences searched: 6
Number of scores above cutoff: 6

Results of the initial comparison of FRAG2 (1-8) with:
File : hpvcomplete.seq



PARAMETERS

		K-tuple	4
Similarity matrix	Unitary	Joining penalty	30
Mismatch penalty	1	Window size	8
Gap penalty	1.00		
Gap size penalty	0.33		
Cutoff score	0		
Randomization group	0		

SEARCH STATISTICS

	Mean	Median	Standard Deviation
Scores:	5	7	0.41
	CPU		Total Elapsed
Times:	00:00:00.00		00:00:00.00

Number of residues: 47269
Number of sequences searched: 6
Number of scores above cutoff: 6

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
	**** 2 standard deviations above mean ****					
1. hpu06714	TOIG of: hpu06714 check: 486	7801	6	7	2.45	0
2. pph16	TOIG of: pph16 check: 6074	7904	6	6	2.45	0
3. a12360	TOIG of: a12360 check: 1580	7909	6	6	2.45	0
4. pph31a	TOIG of: pph31a check: 5866	7912	6	6	2.45	0
5. pph11	TOIG of: pph11 check: 3689	7931	6			
	**** 0 standard deviation from mean ****					
6. af131950	TOIG of: af131950 check: 557	7812	5	6	0.00	0

1. FRAG2 (1-8)
hpu06714 TOIG of: hpu06714 check: 4862 from: 1 to: 7801

TOIG of: hpu06714 check: 4862 from: 1 to: 7801

LOCUS HPU06714 7801 bp DNA VRL 04-FEB-1997
DEFINITION Human papillomavirus HPV-1A (3-3), complete genome.
ACCESSION U06714
VERSION U06714.1 GI:458704
KEYWORDS .
SOURCE Human papillomavirus.
ORGANISM Human papillomavirus
Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
Papillomavirus.
REFERENCE 1 (bases 1 to 7801)
AUTHORS Danos,O., Katinka,M. and Yaniv,M.
TITLE Human papillomavirus 1A complete DNA sequence: a novel type of genome organization among Papovaviridae
JOURNAL EMBO J. 1, 231-236 (1982)
MEDLINE 84182467
REFERENCE 2 (bases 1 to 7801)
AUTHORS Meissner,J.
TITLE Complete nucleotide sequencing of an HPV-1a variant and determination of extant errors in the prototype HPV-1a sequence
JOURNAL Virus Genes 9 (2), 189-191 (1995)
MEDLINE 95250312
REFERENCE 3 (bases 1 to 7801)
AUTHORS Meissner,J.D.
TITLE Direct Submission
JOURNAL Submitted (14-FEB-1994) John D. Meissner, Duke University Medical Center, Microbiology, 277 Carl Building, Research Drive, Durham, NC 27710 USA

FEATURES	Location/Qualifiers
source	1. .7801
	/organism="Human papillomavirus"
	/strain="HPV-1A (3-3)"
	/db_xref="taxon:10566"
variation	142
	/replace="a"
conflict	1283
	/citation=[1]
	/replace="a"
conflict	2301
	/citation=[1]
	/replace="t"
variation	2825
	/replace="t"
conflict	3884. .3886
	/citation=[1]
	/replace="aa"
conflict	4332
	/citation=[1]
	/replace="a"
conflict	4376. .4382
	/citation=[1]
	/replace="gagggaa"
variation	5794
	/replace="a"
variation	6905
	/replace="g"

```
     variation       7186. .7187
                     /note="15 bp deletion"
                     /citation=[1]
     variation       /replace="ctagattgtattgctat"
                     7560
     variation       /replace="c"
                     7618
     conflict        /replace="g"
                     7677. .7678
                     /citation=[1]
     variation       /replace="cc"
                     7787
                     /replace="g"
BASE COUNT      2389 a    1482 c    1664 g    2266 t
ORIGIN

HPU06714  Length: 7801  November 28, 2001 14:10  Type: N  Check: 4862  ..

Initial Score     =      6  Optimized Score  =      7  Significance  =  2.45
Residue Identity  =    87%  Matches          =      7  Mismatches    =     1
Gaps              =      0  Conservative Substitutions               =     0

                                                          X       X
                                                          GACGTTCG
  AAGGAGAATCAACGCCCTCTAGGACACCCGGATCTGTCCCTTCTGCGCGAGACGTTGGAAGTATACACACAA
    3420      3430      3440      3450      3460    X 3470 X    3480

  CGCCTCAAAAGGGACATTCTTCAAGACTTAGACGAC
  3490      3500      3510      3520

2. FRAG2 (1-8)
   pph16          TOIG of: pph16  check: 6074  from: 1  to: 7904

TOIG of: pph16  check: 6074  from: 1  to: 7904

LOCUS       PPH16              7904 bp    DNA   circular  VRL       18-MAR-1994
DEFINITION  Human papillomavirus type 16 (HPV16), complete genome.
ACCESSION   K02718
VERSION     K02718.1  GI:333031
KEYWORDS    circular; complete genome.
SOURCE      Papilloma virus type 16 DNA, isolated from a human invasive
              cervical carcinoma.
  ORGANISM  Human papillomavirus type 16
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 7904)
  AUTHORS   Seedorf,K., Kraemmer,G., Duerst,M., Suhai,S. and Roewekamp,W.G.
  TITLE     Human papillomavirus type 16 DNA sequence
  JOURNAL   Virology 145, 181-185 (1985)
  MEDLINE   85246220
REFERENCE   2  (sites)
  AUTHORS   Kennedy,I.M., Haddow,J.K. and Clements,J.B.
  TITLE     A negative element in the human poapillomavirus type 16 genome acts
              at the level of late mRNA stability
  JOURNAL   J. Virol. 65, 2093-2097 (1991)
  MEDLINE   91162763
COMMENT     The sense strand of this double-stranded circular genome is shown,
            with a numbering system matching the first 60 bp of HPVa1, HPV6b
            and BPV1. The annotation of sites and features is solely based upon
            homology comparison with these other papillomaviruses. In addition
            to the coding sequences reported below, the authors note open
            reading frames which do not start with 'ATG', but which are found
            in other papillomaviruses. In particular, a second portion of the
            E1 gene may be located out to base 2813 (the E1 protein is thought
            to be generally involved in DNA replication).
            A potential 'CAT'-box region is found beginning at base 7895 below,
            and 'TATA' boxes for early and late transcripts may be located at
            17, 65 and 4289. Potential polyadenylation signals are at bases
            4213 and 7260.
            HPV16, in comparison to HPV types 6 and 11, is more often
            associated with malignant genital cancers in humans.

FEATURES             Location/Qualifiers
     source          1. .7904
                     /organism="Human papillomavirus type 16"
                     /db_xref="taxon:10581"
     TATA_signal     17. .23
     TATA_signal     65. .71
     gene            83. .559
                     /gene="E6"
     CDS             83. .559
                     /gene="E6"
                     /note="E6 ORF from 65 to 559; putative"
                     /codon_start=1
                     /product="transforming protein"
                     /protein_id="AAA46939.1"
                     /db_xref="GI:333032"
                     /translation="MHQKRTAMFQDPQERPRKLPQLCTELQTTIHDIILECVYCKQQL
                     LRREVYDFAFRDLCIVYRDGNPYAVCDKCLKFYSKISEYRHYCYSLYGTTLEQQYNKP
                     LCDLLIRCINCQKPLCPEEKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL"
     gene            562. .858
     CDS             /gene="E7"
                     562. .858
                     /gene="E7"
                     /note="E7 ORF from 544 to 858; putative"
                     /codon_start=1
                     /product="transforming protein"
                     /protein_id="AAA46940.1"
                     /db_xref="GI:333033"
                     /translation="MHGDTPTLHEYMLDLQPETTDLYCYEQLNDSSEEEDEIDGPAGQ
     gene            AEPDRAHYNIVTFCCKCDSTLRLCVQSTHVDIRTLEDLLMGTLGIVCPICSQKP"
     CDS             join(865. .1140,1140. .2813)
                     /gene="E1"
                     join(865. .1140,1140. .2813)
                     /gene="E1"
                     /note="E1 interrupted ORF from 859 to 2813; putative"
                     /codon_start=1
                     /product="replication protein"
                     /protein_id="AAA46936.1"
                     /db_xref="GI:459912"
                     /translation="MADPAGTNGEEGTGCNGWFYVEAVVEKKTGDAISDDENENDSDT
                     GEDLVDFIVNDNDYLTQAETETAHALFTAQEAKQHRDAVQVLKRKYLVSPLSDISGCV
                     DNNISPRLKAICIEKQSRAAKRRLFESEDSGYGNTEVETQQMLQVEGRHETETPCSQY
                     SGGSGGGCSQYSSGSGGEGVSERHTICQTPLTNILNVLKTSNAKAAMLAKFKELYGVS
                     FSELVRPFKSNKSTCCDWCIAAFGLTPSIADSIKTLLQQYCLYLHIQSLACSWGMVVL
                     LLVRYKCGKNRETIEKLLSKLLCVSPMCMMIEPPKLRSTAAALYWYKTGISNISEVYG
                     DTPEWIQRQTVLQHSFNDCTFELSQMVQWAYDNDIVDDSEIAYKYAQLADTNSNASAF
                     LKSNSQAKIVKDCATMCRHYKRAEKKQMSMSQWIKYRCDRVDDGGDWKQIVMFLRYQG
                     VEFMSFLTALKRFLQGIPKKNCILLYGAANTGKSLFGMSLMKFLQGSVICFVNSKSHF
                     WLQPLADAKIGMLDDATVPCWNYIDDNLRNALDGNLVSMDVKHRPLVQLKCPPLLITS
     gene            NINAGTDSRWPYLHNRLVVFTFPNEFPFDENGNPVYELNDKNWKSFFSRTWSRLSLHE
                     DEDKENDGDSLPTFKCVSGQNTNTL"
     CDS             2755. .3852
                     /gene="E2"
                     2755. .3852
                     /gene="E2"
                     /note="E2 ORF from 2725 to 3852; putative"
                     /codon_start=1
                     /product="regulatory protein"
                     /protein_id="AAA46941.1"
                     /db_xref="GI:333035"
                     /translation="METLCQRLNVCQDKILTHYENDSTDLRDHIDYWKHMRLECAIYY
                     KAREMGFKHINHQVVPTLAVSKNKALQAIELQLTLETIYNSQYSNEKWTLQDVSLEVY
                     LTAPTGCIKKHGYTVEVQFDGDICNTMHYTNWTHIYICEEASVTVVEGQVDYYGLYYV
                     HEGIRTYFVQFKDDAEKYSKNKVWEVHAGGQVILCPTSVFSSNEVSSPEIIRQHLANH
                     PAATHTKAVALGTEETQTTIQRPRSEPDTGNPCHTTKLLHRDSVDSAPILTAFNSSHK
     gene            GRINCNSNTTPIVHLKGDANTLKCLRYRFKKHCTLYTAVSSTWHWTGHNVKHKSAIVT
                     LTYDSEWQRDQFLSQVKIPKTITVSTGFMSI"
     CDS             3332. .3619
                     /gene="E4"
                     <3332. .3619
                     /gene="E4"
                     /note="E4 ORF from 3332 to 3619; putative"
                     /codon_start=1
                     /protein_id="AAA46937.1"
```

```
                /db_xref="GI:459913"
                /translation="YYVLHLCLAATKYPLLKLLGSTWPTTPPRPIPKPSPWAPKKHRR
                LSSDQDQSQTPETPATPLSCCTETQWTVLQSSLHLTAHTKDGLTVIVTLHP"
gene            3863. .4099
                /gene="E5"
CDS             <3863. .4099
                /gene="E5"
                /note="E5 ORF from 3863 to 4099; putative"
                /codon_start=1
                /protein_id="AAA46938.1"
                /db_xref="GI:459914"
                /translation="YCIHNITGVLFALLCVLLCVCLLIRPLLLSVSTYTSLIILVLLL
                WITAASAFRCFIVYIIFVYIPLFLIHTHARFLIT"
polyA_signal    4213. .4218
                /note="putative"
gene            4235. .5656
                /gene="L2"
CDS             4235. .5656
                /gene="L2"
                /note="L2 ORF from 4133 to 5656; putative"
                /codon_start=1
                /product="minor capsid protein"
                /protein_id="AAA46942.1"
                /db_xref="GI:333036"
                /translation="MRHKRSAKRTKRASATQLYKTCKQAGTCPPDIIPKVEGKTIAEQ
                ILQYGSMGVFFGGLGIGTGSGTGGRTGYIPLGTRPPTATDTLAPVRPPLTVDPVGPSD
                PSIVSLVEETSFIDAGAPTSVPSIPPDVSGFSITTSTDTTPAILDINNTVTTVTTHNN
                PTFTDPSVLQPPTPAETGGHFTLSSSTISTHNYEEIPMDTFIVSTNPNTVTSSTPIPG
                SRPVARLGLYSRTTQQVKVVDPAFVTTPTKLITYDNPAYEGIDVDNTLYFSSNDNSIN
                IAPDPDFLDIVALHRPALTSRRTGIRYSRIGNKQTLRTRSGKSIGAKVHYYYDLSTID
                PAEEIELQTITPSTYTTTSHAASPTSINNGLYDIYADDFITDTSTTPVPSVPSTSLSG
                YIPANTTIPFFGGAYNIPLVSGPDIPINITDQAPSLIPIVPGSPQYTIIADAGDFYLHP
                SYYMLRKRRKRLPYFFSDVSLAA"
TATA_signal     4289. .4295
                /gene="L2"
gene            5559. .7154
                /gene="L1"
CDS             5559. .7154
                /gene="L1"
                /note="L1 ORF from 5526 to 7154; putative"
                /codon_start=1
                /product="major capsid protein"
                /protein_id="AAA46943.1"
                /db_xref="GI:333037"
                /translation="MQVTFIYILVITCYENDVNVYHIFFQMSLWLPSEATVYLPPVPV
                SKVVSTDEYVARTNIYYHAGTSRLLAVGHPYFPIKKPNNNKILVPKVSGLQYRVFRIH
                LPDPNKFGFPDTSFYNPDTQRLVWACVGVEVGRGQPLGVGISGHPLLNKLDDTENASA
                YAANAGVDNRECISMDYKQTQLCLIGCKPPIGEHWGKGSPCTNVAVNPGDCPPLELIN
                TVIQDGDMVHTGFGAMDFTTLQANKSEVPLDICTSICKYPDYIKMVSEPYGDSLFFYL
                RREQMFVRHLFNRAGTVGENVPDDLYIKGSGSTANLASSNYFPTPSGSMVTSDAQIFN
                KPYWLQRAQGHNNGICWGNQLFVTVVDTTRSTNMSLCAAISTSETTYKNTNFKEYLRH
                GEEYDLQFIFQLCKITLTADVMTYIHSMNSTILEDWNFGLQPPPGGTLEDTYRFVTQA
                IACQKHTPPAPKEDDPLKKYTFWEVNLKEKFSADLDQFPLGRKFLLQAGLKAKPKFTL
                GKRKATPTTSSTSTTAKRKKRKL"
    polyA_signal    7260. .7265
BASE COUNT      2601 a   1377 c   1509 g   2417 t
ORIGIN      Unreported.

  PPH16  Length: 7904  November 28, 2001 14:10  Type: N  Check: 6074  ..

Initial Score   =      6  Optimized Score  =      6  Significance  =   2.45
Residue Identity =    75%  Matches          =      6  Mismatches    =      2
Gaps            =      0  Conservative Substitutions                =      0

                                                      X      X
                                                      GACGTTCG
                                                      ||||||
     TTTGTTTGTTTGTTTTTTAATAAACTGTTATTACTTAACAATGCGACACAAACGTTCTGCAAAACGCACAAA
       4200      4210      4220      4230      4240    X 4250 X    4260

     ACGTGCATCGGCTACCCAACTTTATAAAACATGCAA
       4270      4280      4290      4300

3. FRAG2 (1-8)
   a12360      TOIG of: a12360  check: 1580  from: 1  to: 7909

 TOIG of: a12360  check: 1580  from: 1  to: 7909

LOCUS       A12360          7909 bp    DNA                PAT       12-DEC-1993
DEFINITION  complete nucleotide sequence of HPV-33.
ACCESSION   A12360
VERSION     A12360.1  GI:492936
KEYWORDS    .
SOURCE      Human papillomavirus type 33.
  ORGANISM  Human papillomavirus type 33
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 7909)
  AUTHORS   .
  JOURNAL   Patent: WO 8705630-A 1 24-SEP-1987;
FEATURES             Location/Qualifiers
     source          1. .7909
                     /organism="Human papillomavirus type 33"
                     /db_xref="taxon:10586"
BASE COUNT     2544 a   1354 c   1535 g   2474 t      2 others
ORIGIN

  A12360  Length: 7909  November 28, 2001 14:10  Type: N  Check: 1580  ..

Initial Score   =      6  Optimized Score  =      6  Significance  =   2.45
Residue Identity =    75%  Matches          =      6  Mismatches    =      2
Gaps            =      0  Conservative Substitutions                =      0

                                                      X      X
                                                      GACGTTCG
                                                      ||||||
     TGTGTGCACTAAAACGAAAGTTTGCCGCATGTTCACAAAGTGCTGCGGAGGACGTTGTTGATCGTGCTGCAA
       1130      1140      1150      1160      1170  X   1180      1190

     ACCCGTGTAGAACGTCTATTAATAAAAATAAAGAAT
       1200      1210      1220      1230

4. FRAG2 (1-8)
   pph31a      TOIG of: pph31a  check: 5866  from: 1  to: 7912

 TOIG of: pph31a  check: 5866  from: 1  to: 7912

LOCUS       PPH31A          7912 bp    DNA   circular  VRL      18-MAR-1994
DEFINITION  Human papillomavirus type 31 (HPV-31) complete genome.
ACCESSION   J04353
VERSION     J04353.1  GI:333048
KEYWORDS    complete genome.
SOURCE      Human papillomavirus type 31 DNA.
  ORGANISM  Human papillomavirus type 31
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 7912)
  AUTHORS   Goldsborough,M.D., DiSelvestre,D., Temple,G.F. and Lorincz,A.T.
  TITLE     Nucleotide sequence of human papillomavirus type 31: A cervical
            neoplasia associated virus
  JOURNAL   Virology 171, 306-311 (1989)
  MEDLINE   89299478
COMMENT     Draft entry and computer-readable copy of sequence [1] kindly
            submitted by M.D.Goldsborough, 05-JUL-1989.
FEATURES             Location/Qualifiers
     source          1. .7912
                     /organism="Human papillomavirus type 31"
                     /db_xref="taxon:10585"
     TATA_signal     19. .24
     TATA_signal     69. .74
     gene            108. .557
                     /gene="E6"
     CDS             108. .557
                     /gene="E6"
```

```
                     /note="ORF E6 from bp 39 to 557"
                     /codon_start=1
                     /product="transforming protein"
                     /protein_id="AAA46950.1"
                     /db_xref="GI:459916"
                     /translation="MFKNPAERPRKLHELSSALEIPYDELRLNCVYCKGQLTETEVLD
                     FAFTDLTIVYRDDTPHGVCTKCLRFYSKVSEFRWYRYSVYGTTLEKLTNKGICDLLIR
misc_feature         CITCQRPLCPEEKQRHLDKKKRFHNIGGRWTGRCIACWRRPRTETQV"
                     228. .236
                     /gene="E6"
misc_feature         /standard_name="Splice donor"
                     403. .414
                     /gene="E6"
gene                 /standard_name="Splice acceptor"
                     560. .856
CDS                  /gene="E7"
                     560. .856
                     /gene="E7"
                     /note="ORF E7 from bp 545 to 856"
                     /codon_start=1
                     /product="transforming protein"
                     /protein_id="AAA46951.1"
                     /db_xref="GI:459917"
                     /translation="MRGETPTLQDYVLDLQPEATDLHCYEQLPDSSDEEDVIDSPAGQ
gene                 AEPDTSNYNIVTFCCQCKSTLRLCVQSTQVDIRILQELLMGSFGIVCPNCSTRL"
                     862. .2751
CDS                  /gene="E1"
                     862. .2751
                     /gene="E1"
                     /note="ORF E1 from bp 850 to bp 2751"
                     /codon_start=1
                     /product="replication protein"
                     /protein_id="AAA46952.1"
                     /db_xref="GI:459918"
                     /translation="MADPAGTDGEGTGCNGWFYVEAVIDRQTGDNISEDENEDSSDTG
                     EDMVDFIDNCNVYNNQAEAETAQALFHAQEAEEHAEAVQVLKRKYVGSPLSDISSCVD
                     YNISPRLKAICIENNSKTAKRRLFELPDSGYGNTEVETQQMVQVEEQQTTLSCNGSDG
                     THSERENETPTRNILQVLKTSNGKAAMLGKFKELYGVSFMELIRPFQSNKSTCTDWCV
                     AAFGVTGTVAEGFKTLLQPYCLYCHLQSLACSWGMVMLMLVRFKCAKNRITIEKLLEK
                     LLCISTNCMLIQPPKLRSTAAALYWYRTGMSNISDVYGETPEWIERQTVLQHSFNDTT
                     FDLSQMVQWAYDNDVMDDSEIAYKYAQLADSDSNACAFLKSNSQAKIVKDCGTMCRHY
                     KRAEKRQMSMGQWIKSRCDKVSDEGDWRDIVKFLRYQQIEFVSFLSALKLFLKGVPKK
                     NCILIHGAPNTGKSYFGMSLISFLQGCIISYANSKSHFWLQPLADAKIGMLDDATTPC
                     WHYIDNYLRNALDGNPVSIDVKHKALMQLKCPPLLITSNINAGKDDRWPYLHSRLVVF
                     TFPNPFPFDKNGNPVYELSDKNWKSFFSRTWCRLNLHEEEDKENDGDSFSTFKCVSGQ
gene                 NIRTL"
                     2693. .3811
CDS                  /gene="E2"
                     2693. .3811
                     /gene="E2"
                     /note="ORF E2 from bp 2663 to 3811"
                     /codon_start=1
                     /product="regulatory protein"
                     /protein_id="AAA46953.1"
                     /db_xref="GI:459919"
                     /translation="METLSQRLNVCQDKILEHYENDSKRLCDHIDYWKHIRLECVLMY
                     KAREMGIHSINHQVVPALSVSKAKALQAIELQMMLETLNNTEYKNEDWTMQQTSLELY
                     LTAPTGCLKKHGYTVEVQFDGDVHNTMHYTNWKFIYLCIDGQCTVVEGQVNCKGIYYV
                     HEGHITYFVNFTEEAKKYGTGKKWEVHAGGQVIVFPESVFSSDEISFAGIVTKLPTAN
                     NTTTSNSKTCALGTSEGVRRATTSTKRPRTEPEHRNTHHPNKLLRGDSVDSVNCGVIS
                     AAACTNQTRAVSCPATTPIIHLKGDANILKCLRYRLSKYKQLYEQVSSTWHWTCTDGK
gene                 HKNAIVTLTYISTSQRDDFLNTVKIPNTVSVSTGYMTI"
                     3270. .3578
CDS                  /gene="E4"
                     <3270. .3578
                     /gene="E4"
                     /note="ORF E4 from bp 3270 to 3578"
                     /codon_start=1
                     /protein_id="AAA46949.1"
                     /db_xref="GI:459915"
                     /translation="LFFLNLYLAVTKYPLLGLLQSYQQPTTPPHRIPKPAPWAPVKVC
                     GGRRRLLSDQEQSQSTETPTTPTSCCEATPWTVSTVGLSVQLHAQTKQGLSVVLQLHL
                     "
```

```
gene                 3816. .4070
CDS                  /gene="E5"
                     3816. .4070
                     /gene="E5"
                     /note="ORF E5 from bp 3804 to 4070"
                     /codon_start=1
                     /protein_id="AAA46954.1"
                     /db_xref="GI:459920"
                     /translation="MIELNISTVSIVLCFLLCFCVLLFVCLVIRPLVLSVSVYATLLL
repeat_region        LIVILWVIATSPLRCFCIYVVFIYIPLFVIHTHASFLSQQ"
polyA_signal         4099. .4134
                     4138. .4143
repeat_region        /note="putative"
gene                 4143. .4158
                     4171. .5571
CDS                  /gene="L2"
                     4171. .5571
                     /gene="L2"
                     /note="ORF L2 from bp 4060 to bp 5571"
                     /codon_start=1
                     /product="minor capsid protein"
                     /protein_id="AAA46955.1"
                     /db_xref="GI:459921"
                     /translation="MRSKRSTKRTKRASATQLYQTCKAAGTCPSDVIPKIEHTTIADQ
                     ILRYGSMGVFFGGLGIGSGSGTGGRTGYVPLSTRPSTVSEASIPIRPPVSIDPVGPLD
                     PSIVSLVEESGIVDVGAPAPIPHPPTTSGFDIATTADTTPAILDVTSVSTHENPTFTD
                     PSVLQPPTPAETSGHLLLSSSSISTHNYEEIPMDTFIVSTNNENITSSTPIPGVRRPA
                     RLGLYSKATQQVKVIDPTFLSAPKQLITYENPAYETVNAEESLYFSNTSHNIAPDPDF
                     LDIIALHRPALTSRRNTVRYSRLGNKQTLRTRSGATIGARVHYYYDISSINPAGESIE
                     MQPLGASATTTSTLNDGLYDIYADTDFTVDTPATHNVSPSTAVQSTSAVSAYVPTNTT
                     VPLSTGFDIPIFSGPDVPIEHAPTQVFPFPLAPTTPQVSIFVDGGDFYLHPSYYMLKR
gene                 RRKRVSYFFTDVSVAA"
                     5552. .7066
CDS                  /gene="L1"
                     5552. .7066
                     /gene="L1"
                     /note="ORF L1 from bp 5516 to  7066"
                     /codon_start=1
                     /product="major capsid protein"
                     /protein_id="AAA46956.1"
                     /db_xref="GI:459922"
                     /translation="MSLWRPSEATVYLPPVPVSKVVSTDEYVTRTNIYYHAGSARLLT
                     VGHPYYSIPKSDNPKKIVVPKVSGLQYRVFRVRLPDPNKFGFPDTSFYNPETQRLVWA
                     CVGLEVGRGQPLGVGISGHPLLNKFDDTENSNRYAGGPGTDNRECISMDYKQTQLCLL
                     GCKPPIGEHWGKGSPCSNNAITPGDCPPLELKNSVIQDGDMVDTGFGAMDFTALQDTK
                     SNVPLDICNSICKYPDYLKMVAEPYGDTLFFYLRREQMFVRHFFNRSGTVGESVPTDL
                     YIKGSGSTATLANSTYFPTPSGSMVTSDAQIFNKPYWMQRAQGHNNGICWGNQLFVTV
                     VDTTRSTNMSVCAAIANSDTTFKSSNFKEYLRHGEEFDLQFIFQLCKITLSADIMTYI
                     HSMNPAILEDWNFGLTTPPSGSLEDTYRFVTSQAITCQKTAPQKPKEDPFKDYVFWEV
                     NLKEKFSADLDQFPLGRKFLLQAGYRARPKFKAGKRSAPSASTTTPAKRKKTKK"
polyA_signal         7227. .7231
                     /note="putative"
repeat_region        7291. .7302
repeat_region        7314. .7333
                     /note="putative"
protein_bind         7406. .7420
                     /standard_name="glucocorticoid responsive element"
                     /bound_moiety="hormone receptor"
protein_bind         7477. .7488
                     /function="gene transcription"
                     /bound_moiety="E2"
enhancer             7542. .7549
                     /standard_name="keratinocyte-dependant enhancer"
protein_bind         7868. .7879
                     /function="gene transcription"
                     /bound_moiety="E2"
BASE COUNT       2528 a    1364 c    1572 g    2448 t
ORIGIN

PPH31A  Length: 7912  November 28, 2001 14:10  Type: N  Check: 5866  ..

Initial Score   =      6  Optimized Score  =      6  Significance =  2.45
Residue Identity =    75%  Matches         =      6  Mismatches   =     2
```

```
Gaps            =      0  Conservative Substitutions                =      0

                                                          X      X
                                                          GACGTTCG
                                                          ||||||
   ACATTTGGATAAAAAGAAACGATTCCACAACATAGGAGGAAGGTGGACAGGACGTTGCATAGCATGTTGGAG
   460       470       480       490       500       510     X 520

   AAGACCTCGTACTGAAACCCAAGTGTAAACATGCGT
 530       540       550       560


5  FRAG2 (1-8)
   pph11          TOIG of: pph11  check: 3689  from: 1  to: 7931

   TOIG of: pph11  check: 3689  from: 1  to: 7931

LOCUS       PPH11          7931 bp    DNA   circular  VRL       02-JUN-1994
DEFINITION  Human papillomavirus type 11 (HPV-11) complete genome.
ACCESSION   M14119
VERSION     M14119.1  GI:333026
KEYWORDS    complete genome.
SOURCE      Human laryngeal papillomavirus type 11 DNA.
  ORGANISM  Human papillomavirus type 11
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 7931)
  AUTHORS   Dartmann,K., Schwarz,E., Gissmann,L. and Zur Hausen,H.
  TITLE     The nucleotide sequence and genome organization of human papilloma
            virus type 11
  JOURNAL   Virology 151, 124-130 (1986)
  MEDLINE   86181601
COMMENT     ORF L1 is assumed to encode the major structural protein.
FEATURES             Location/Qualifiers
     source          1. .7931
                     /organism="Human papillomavirus type 11"
                     /db_xref="taxon:10580"
     CAAT_signal     9. .15
                     /note="putative"
     protein_bind    35. .46
                     /note="putative"
                     /function="gene transcription"
                     /bound_moiety="E2"
     protein_bind    50. .61
                     /note="putative"
                     /function="gene transcription"
                     /bound_moiety="E2"
     TATA_signal     66. .71
                     /note="putative"
     gene            102. .554
                     /gene="E6"
     CDS             102. .554
                     /gene="E6"
                     /note="102 is position of first start codon in ORF E6;
                     putative"
                     /codon_start=1
                     /product="transforming protein"
                     /protein_id="AAA46927.1"
                     /db_xref="GI:496193"
                     /translation="MESKDASTSATSIDQLCKTFNLSLHTLQIQCVFCRNALTTAEIY
                     AYAYKNLKVVWRDNFPFAACACCLELQGKINQYRHFNYAAYAPTVEEETNEDILKVLI
                     RCYLCHKPLCEIEKLKHILGKARFIKLNNQWKGRCLHCWTTCMEDLLP"
     gene            530. .826
                     /gene="E7"
     CDS             530. .826
                     /gene="E7"
                     /note="530 is position of first start codon in ORF E7;
                     putative"
                     /codon_start=1
                     /product="transforming protein"
                     /protein_id="AAA46928.1"
                     /db_xref="GI:496194"
                     /translation="MHGRLVTLKDIVLDLQPPDPVGLHCYEQLEDSSEDEVDKVDKQD
                     AQPLTQHYQILTCCCGCDSNVRLVVECTDGDIRQLQDLLLGTLNIVCPICAPKP"
     gene            832. .2781
                     /gene="E1"
     CDS             832. .2781
                     /gene="E1"
                     /note="832 is position of first start codon in ORF E1;
                     putative"
                     /codon_start=1
                     /product="replication protein"
                     /protein_id="AAA46929.1"
                     /db_xref="GI:496195"
                     /translation="MADDSGTENEGSGCTGWFMVEAIVEHTTGTQISEDEEEEVEDSG
                     YDMVDFIDDRHITQNSVEAQALFNRQEADAHYATVQDLKRKYLGSPYVSPISNVANAV
                     ESEISPRLDAIKLTTQPKKVKRRLFETRELTDSGYGYSEVEAATQVEKHGDPENGGDG
                     QERDTGRDIEGEGVEHREAEAVDDSTREHADTSGILELLKCKDIRSTLHGKFKDCFGL
                     SFVDLIRPFKSDRTTCADWVVAGFGIHHSIADAFQKLIEPLSLYAHIQWLTNAWGMVL
                     LVLIRFKVNKSRCTVARTLGTLLNIPENHMLIEPPKIQSGVRALYWFRTGISNASTVI
                     GEAPEWITRQTVIEHSLADSQFKLTEMVQWAYDNDICEESEIAFEYAQRGDFDSNARA
                     FLNSNMQAKYVKDCAIMCRHYKHAEMKKMSIKQWIKYRGTKVDSVGNWKPIVQFLRHQ
                     NIEFIPFLSKLKLWLHGTPKKNCIAIVGPPDTGKSCFCMSLIKFLGGTVISYVNSCSH
                     FWLQPLTDAKVALLDDATQPCWTYMDTYMRNLLDGNPMSIDRKHRALTLIKCPPLLVT
                     SNIDISKEEKYKYLHSRVTTFTFPNPFPFDRNGNAVYELSDANWKCFFERLSSSLDIE
                     DSEDEEDGSNSQAFRCVPGSVVRTL"
     gene            2723. .3826
                     /gene="E2"
     CDS             2723. .3826
                     /gene="E2"
                     /note="2723 is position of first start codon in ORF E2;
                     putative"
                     /codon_start=1
                     /product="regulatory protein"
                     /protein_id="AAA46930.1"
                     /db_xref="GI:496196"
                     /translation="MEAIAKRLDACQDQLLELYEENSIDIHKHIMHWKCIRLESVLLH
                     KAKQMGLSHIGLQVVPPLTVSETKGHNAIEMQMHLESLAKTQYGVEPWTLQDTSYEMW
                     LTPPKRCFKKQGNTVEVKFDGCEDNVMEYVVWTHIYLQDNDSWVKVTSSVDAKGIYYT
                     CGQFKTYYVNFNKEAQKYGSTNHWEVCYGSTVICSPASVSSTVREVSIAEPTTYTPAQ
                     TTAPTVSACTTEDGVSAPPRKRARGPSTNNTLCVANIRSVDSTINNIVTDNYNKHQRR
                     NNCHSAATPIVQLQGDSNCLKCFRYRLNDKYKHLFELASSTWHWASPEAPHKNAIVTL
                     TYSSEEQRQQFLNSVKIPPTIRHKVGFMSLHLL"
     gene            3255. .3581
                     /gene="E4"
     CDS             3255. .3581
                     /gene="E4"
                     /note="3255 is position of first start codon in ORF E4;
                     putative"
                     /codon_start=1
                     /protein_id="AAA46931.1"
                     /db_xref="GI:496197"
                     /translation="MVVPIIGKYVMAAQLYVLLHLYLALYEKYPLLNLLHTPPHRPPP
                     LQCPPAPRKTACRRRLGSEHVDRPLTTPCVWPTSDPWTVQSTTSSLTITTSTKEGTTV
                     TVQLRL"
     gene            3871. .4146
                     /gene="E5A"
     CDS             3871. .4146
                     /gene="E5A"
                     /note="3871 is position of first start codon in ORFE5A;
                     putative"
                     /codon_start=1
                     /protein_id="AAA46932.1"
                     /db_xref="GI:496198"
                     /translation="MEVVPVQIAAATTTTLILPVVIAFAVCILSIVLIILISDFVVYT
                     SVLVLTLLLYLLLWLLLTTPLQFFLLTLCVCYFPAFYIHIYIVQTQQ"
     gene            4146. .4370
                     /gene="E5B"
     CDS             4146. .4370
                     /gene="E5B"
                     /note="4146 is position of first start codon in ORF E5B;
                     putative"
                     /codon_start=1
                     /protein_id="AAA46933.1"
                     /db_xref="GI:496199"
                     /translation="MVMLTCHLNDGDTWLFLWLFTAFVVAVLGLLLLHYRAVHGTEKT
```

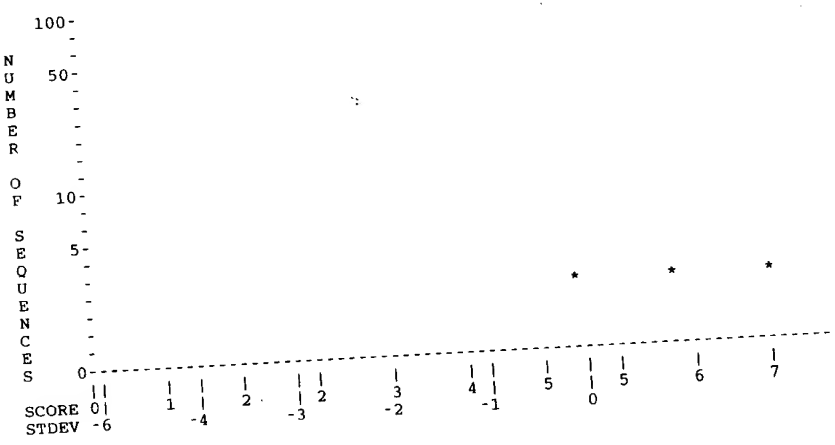
```
     polyA_signal     KCAKCKSNRNTTVDYVYMSHGDNGDYVYMN"
                      4371. .4376
     gene             /note="putative"
                      4417. .5784
     CDS              /gene="L2"
                      4417. .5784
                      /gene="L2"
                      /note="4417 is position of first start codon in ORF L2;
                      putative"
                      /codon_start=1
                      /product="minor capsid protein"
                      /protein_id="AAA46934.1"
                      /db_xref="GI:496200"
                      /translation="MKPRARRRKRASATQLYQTCKATGTCPPDVIPKVEHTTIADQIL
                      KWGSLGVFFGGLGIGTGAGSGGRAGYIPLGSSPKPAITGGPAARPPVLVEPVAPSDPS
                      IVSLIEESAIINAGAPEVVPPTQGGFTITSSESTTPAILDVSVTNHTTTSVFQNPLFT
                      EPSVIQPQPPVEASGHILISAPTITSQHVEDIPLDTFVVSSSDSGPTSSTPLPRAFPR
                      PRVGLYSRALQQVQVTDPAFLSTPQRLVTYDNPVYEGEDVSLQFTHESIHNAPDEAFM
                      DIIRLHRPAITSRRGLVRFSRIGQRGSMYTRSGQHIGARIHYFQDISPVTQAAEEIEL
                      HPLVAAENDTFDIYAEPFDPIPDPVQHSVTQSYLTSTPNTLSQSWGNTTVPLSIPSDW
                      FVQSGPDITFPTASMGTPFSPVTPALPTGPVFITGSDFYLHPTWYFARRRRKRIPLFF
                      TDVAA"
     polyA_signal     4545. .4551
                      /gene="L2"
     gene             /note="putative"
                      5771. .7276
     CDS              /gene="L1"
                      5771. .7276
                      /gene="L1"
                      /note="5771 is position of first start codon in ORF L1;
                      putative"
                      /codon_start=1
                      /product="major capsid protein"
                      /protein_id="AAA46935.1"
                      /db_xref="GI:496201"
                      /translation="MWRPSDSTVYVPPPNPVSKVVATDAYVKRTNIFYHASSSRLLAV
                      GHPYYSIKKVNKTVVPKVSGYQYRVFKVVLPDPNKFALPDSSLFDPTTQRLVWACTGL
                      EVGRGQPLGVGVSGHPLLNKYDDVENSGGYGGNPGQDNRVNVGMDYKQTQLCMVGCAP
                      PLGEHWGKGTQCSNTSVQNGDCPPLELITSVIQDGDMVDTGFGAMNFADLQTNKSDVP
                      LDICGTVCKYPDYLQMAADPYGDRLFFYLRKEQMFARHFFNRAGTVGEPVPDDLLVKG
                      GNNRSSVASSIYVHTPSGSLVSSEAQLFNKPYWLQKAQGHNNGICWGNHLFVTVVDTT
                      RSTNMTLCASVSKSATYTNSDYKEYMRHVEEFDLQFIFQLCSITLSAEVMAYIHTMNP
                      SVLEDWNFGLSPPPNGTLEDTYRYVQSQAITCQKPTPEKEKQDPYKDMSFWEVNLKEK
     5'UTR            FSSELDQFPLGRKFLLQSGYRGRTSARTGIKRPAVSKPSTAPKRKRTKTKK"
                      join(7277. .7931,1. .101)
                      /standard_name="LCR"
                      /function="regulatory region"
     repeat_region    7320. .7360
                      /standard_name="direct repeats"
                      /note="putative"
     repeat_region    7339. .7374
                      /standard_name="direct repeats"
                      /note="putative"
     repeat_region    7374. .7403
     polyA_signal     7457. .7462
                      /note="putative"
     protein_bind     7592. .7603
                      /note="putative"
                      /function="gene transcription"
                      /bound_moiety="E2"
     polyA_signal     7748. .7753
                      /note="putative"
     protein_bind     7890. .7901
                      /note="putative"
                      /function="gene transcription"
                      /bound_moiety="E2"
BASE COUNT      2406 a    1519 c    1736 g    2270 t
ORIGIN      4557 bp upstream of HindIII site.

  PPH11  Length: 7931  November 28, 2001 14:10  Type: N  Check: 3689  ..

Initial Score  =       6  Optimized Score  =       6  Significance =  2.45

Residue Identity =     75%  Matches          =       6  Mismatches   =      2
Gaps             =       0  Conservative Substitutions               =      0

                                                          X       X
                                                          GACGTTCG
                                                          ||||||
      GAAAGTAAAGATGCCTCCACGTCTGCAACATCTATAGACCAGTTGTGCAAGACGTTTAATCTTTCTTTGCAC
        110       120       130       140       150     X  160 X      170
      ACTCTGCAAATTCAGTGCGTGTTTTGCAGGAATGCA
       180       190       200       210

6. FRAG2 (1-8)
   af131950      TOIG of: af131950  check: 557  from: 1  to: 7812

  TOIG of: af131950  check: 557  from: 1  to: 7812

LOCUS       AF131950      7812 bp    DNA                 VRL       07-FEB-2001
DEFINITION  Human papillomavirus candHPV85, complete genome.
ACCESSION   AF131950
VERSION     AF131950.1  GI:4574720
KEYWORDS    .
SOURCE      Human papillomavirus candHPV85.
  ORGANISM  Human papillomavirus candHPV85
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 7812)
  AUTHORS   Chow,V.T.K. and Leong,P.W.F.
  TITLE     Complete nucleotide sequence, genomic organization and phylogenetic
            analysis of a novel genital human papillomavirus type, HLT7474-S
  JOURNAL   J. Gen. Virol. 80 (Pt 11), 2923-2929 (1999)
  MEDLINE   20047972
   PUBMED   10580054
REFERENCE   2  (bases 1 to 7812)
  AUTHORS   Chow,V.T.K. and Leong,W.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-FEB-1999) Department of Microbiology, National
            University of Singapore, 10 Kent Ridge Crescent, Singapore 119260,
            Singapore
FEATURES             Location/Qualifiers
     source          1. .7812
                     /organism="Human papillomavirus candHPV85"
                     /db_xref="taxon:151757"
                     /note="isolated from scraped uterine cervical cells from a
     protein_bind    female sex worker; overlapping PCR products"
                     40. .51
                     /note="putative"
                     /bound_moiety="E2 protein"
     protein_bind    /function="transcriptional regulation"
                     56. .67
                     /note="putative"
                     /bound_moiety="E2 protein"
     TATA_signal     /function="transcriptional regulation"
                     71. .79
     gene            /note="putative"
                     105. .578
     CDS             /gene="E6"
                     105. .578
                     /gene="E6"
                     /codon_start=1
                     /product="putative transforming protein E6"
                     /protein_id="AAD24181.1"
                     /db_xref="GI:4574721"
                     /translation="MAEFGNPATRPYKLPDLCNTLDTSLQDIEISCVYCKSVLQRTEV
                     YEFAFADLFVVYRDGIPYAACQNCLMFYSKIRELRYYSDSVYGETLEKLTNSNIYDIL
     gene            IRCLRCQKPLCPAEKLKHLNEKRRFHKIAGKYRGQCRRCMTRAQEQQGSRRETQV"
                     587. .913
     CDS             /gene="E7"
                     587. .913
                     /gene="E7"
                     /codon_start=1
                     /product="putative transforming protein E7"
```

```
                     /protein_id="AAD24182.1"
                     /db_xref="GI:4574722"
                     /translation="MHGPKPTVHDIVLDLEPYNEVQEVDLYCYEELNNSEEEIDEPDN
                     AINHRQPLLARREELQRHTICCVCCKCEASLQLVVESSAADLRDLQQLFLGTLSFLCP
                     LCAVLR"
     gene            920. .2866
                     /gene="E1"
     CDS             920. .2866
                     /gene="E1"
                     /codon_start=1
                     /product="putative replication protein E1"
                     /protein_id="AAD24183.1"
                     /db_xref="GI:4574723"
                     /translation="MADPEGTDGDGTGCNGWFFVQAIVDKKTGDTISEDEDEDGTDTG
                     SDLVDFIDTSNTNYMQADREAAQALLHAQEVETDTKLLHALKRKYGAHSTENSPCRDT
                     ASIHSNLSSPLQEISLNSYNNTAKRRLCSVPDSGYGNTQVETLQTQVTLDTNVFGDGK
                     NGDGLNSEACSTDNEMDIENQNPNSPMTQIVSLLKVNNKKAAILAKFKETYGLSFTDL
                     VRTFKSDKTTCTDWVAAICGVNPNIAEGFKTLIQPYVLYAHIQCMDCSWGVFILALLR
                     YKCGKNRLTVAKGLSTLLHVPDTHMLIEPPKLRSSCAALYWYRTGISNISEVTGDTPE
                     WIQRQTIIQHGIDDSVFDLSEMIQWAFDNDYIDESDIAYEYAQLADCNSNAAAFLKSN
                     CQAKYLRDCAVMCRHYKRAQRKQMNMSQWISYRCDKIDDGGDWKPIVQFLRFQGIEFI
                     TFLRAFKDFLKGTPKKNCIVIYGPANTGKSYFCMSLIQFLHGTVLSFVNSNSHFWLEP
                     LTDTKIAMVDDATPTCWSYFDNYMRNALDGNPISIDRKHKHLIQMKCPPMLITSNTNP
                     ATDDRWPYLRSRVTVFTFPHTFPFDSNGNPVYDINDKNWKCFFKRTWSRLDLHQEEED
                     TENDGCPIPTFKCVTGKDSRTLRD"
     gene            2799. .3932
                     /gene="E2"
     CDS             2799. .3932
                     /gene="E2"
                     /codon_start=1
                     /product="putative regulatory protein E2"
                     /protein_id="AAD24184.1"
                     /db_xref="GI:4574724"
                     /translation="MMDALSQRLNVLQEKILEHYETDSTDICDQIDYWKCVRLENAIY
                     YAAREHGLKTINHHVVPTFQISKSKAHEAIELQMALESLAKSEFKNELWTLQDTCQEL
                     YQTPPQQCFKKQGQTVEVRYDGDKDNTMHYTSWDYIYYVTEGDKWCKTKGYVNYCGLY
                     YIKEGQQTYYVQFKCDAQQYGQSGKWEVWYNGKKIECVNTCDKIIQCSESMYSTCDET
                     VSATAIARELQHPTTPYTEATTVCTQKSGGSAPTRNPFRHCGFTETSEVDGLSVDHLN
                     NPLLSAPAGNNGFRKHNSGDTTPIVHLKGDKNRLKCLRYRLQKFHNLYTQISCTWHWI
                     GGKETTRTGILTLAYANETQRQKFLDVVKIPNTVNVSLGYMTL"
     gene            <3412. .3699
                     /gene="E4"
     CDS             <3412. .3699
                     /gene="E4"
                     /note="start codon absent"
                     /codon_start=1
                     /product="putative protein E4"
                     /protein_id="AAD24185.1"
                     /db_xref="GI:4574725"
                     /translation="YNVLNLCTVPVTRQYPLLQLLENYNTPPHRIPKPPPCAPKRAGV
                     RRRLETHSDIVDSQKQVKSTDCPWTTSTTPCLVHLQATTGSGSTIVVTLRL"
     gene            3940. .4161
                     /gene="E5"
     CDS             3940. .4161
                     /gene="E5"
                     /codon_start=1
                     /product="putative protein E5"
                     /protein_id="AAD24186.1"
                     /db_xref="GI:4574726"
                     /translation="MLIAVFLLCFCVCLYVCCTIPLFPATNVCAHAWLLVFLLIVFIT
                     TPFTAFIVYIFFFILPMFLLHLHSVHTFD"
     gene            4226. .5629
                     /gene="L2"
     CDS             4226. .5629
                     /gene="L2"
                     /codon_start=1
                     /product="putative minor capsid protein L2"
                     /protein_id="AAD24187.1"
                     /db_xref="GI:4574727"
                     /translation="MVSHRAARRKRASATDLYKTCKQSGTCPPDVINKVEGTTLADKL
                     LQWTSLGIFLGGLGIGTGSGTGGRTGYIPLGGRPNTVVDVSPARPPVVIESVGPSDPS
                     IVTLVEESSIVTSGAPVPTFTGTSGFEITSSATTTPAVLDITPASGSVQLSSTSFTNP
                     AFTDPSVIEVPQTGEVSGDIFITTPTSGTHGYEEIPMHTFATQGRGTEPISSTPIPGV
                     RRVAGPRLYSQAYQQVKITNSDFISRPSTLVTFTNPAYEPIDTTLTFSPQDVVPDPDF
                     MDIVRLHRPALTSRRGTVRFSRLGKKLTMSTRSGKQIGAQVHYYHDISPISHIGESIE
                     MQPLLPDAAVTADTNGLFDIYADTDIDNNAMLYDRNISDVTQPTTSTISSVSSRYSNT
                     TIPLATSWDVPVHTGPDMTLPTTIPQWPNIVPLLPNNTHSVVLQGTNYYLWPNYYFIF
                     KKKRKRVPYFLTDGFVAF"
     gene            5610. .7124
                     /gene="L1"
     CDS             5610. .7124
                     /gene="L1"
                     /codon_start=1
                     /product="putative major capsid protein L1"
                     /protein_id="AAD24188.1"
                     /db_xref="GI:4574728"
                     /translation="MALWRSSDSKVYLPPPSVAKVVNTDDYVTRTSTFYHAGSSRLLT
                     VGHPYYKVTSNGGRKQDIPKVSAYQYRVFRVTLPDPNKFGLPDTNIYNPETQRLVWAC
                     VGMEVGRGQPLGVSLSGHPFYNKLDDTENSHVATSVVTHDTRDNVSVDYKQTQLCIIG
                     CVPAIGEHWAKGTACKPGAVQTGDCPPLELVNTPIEDGDMIDTGYGAMDFSTLQDNKS
                     EVPLDICQSICKYPDYLQMSADAYGDSMFFCLRREQLFARHFWNRGGTIGDAVPETLY
                     IKGTNDRATPGSCIYSPSPSGSMVSSDAQMFNKPYWLHKAQGQNNGICWHNQLFITVV
                     DTTRSTNLTLSTATTNPVPSIYEPSKFKEYTRHVEEYDLQFIFQLCKITLTTDVMSYI
                     HNMDPTILDSWNFGVSPPPSASLVDTYRFLQSSAITCQKDVVVPQKKDPYEKLKFWNV
                     DLKEHFSSDLDQFPLGRKFLLQAGLRPKPTIGPRKRVASTSTATRPSKRKRTAK"
     protein_bind    7456. .7467
                     /note="putative"
                     /bound_moiety="E2 protein"
                     /function="transcriptional regulation"
     misc_feature    7647. .7666
                     /note="consensus motif of oncogenic anogenital HPV types"
     protein_bind    7776. .7787
                     /note="putative"
                     /bound_moiety="E2 protein"
                     /function="transcriptional regulation"

BASE COUNT      2502 a    1391 c    1558 g    2361 t
ORIGIN

AF131950  Length: 7812  November 28, 2001 14:10  Type: N  Check: 557  ..

Initial Score  =      5  Optimized Score  =      6  Significance  =  0.00
Residue Identity =  75%  Matches          =      6  Mismatches    =     2
Gaps           =      0  Conservative Substitutions               =     0

                                                        X       X
                                                        GACGTTCG
                                                        ||||| |

   TATTCTAAAATAAGAGAACTAAGATATTATAGCGACTCTGTGTATGGGGAAACGTTAGAAAAACTAACAAAT
   320       330       340       350       360       370     X 380

   AGCAATATATATGATATATTAATAAGGTGTTTACGG
   390       400       410       420
```

frag2-inv.res

Mon Dec 3 08:02:33 2001

```
> O <
O| |O  IntelliGenetics
> O <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file frag2-inv.res made by sdavid on Wed 28 Nov 101 14:20:33-PST.

Query sequence being compared:          FRAG2' (1-8)
Number of sequences searched:                      6
Number of scores above cutoff:                     6

      Results of the initial comparison of FRAG2' (1-8) with:
        File : hpvcomplete.seq
```



PARAMETERS

Similarity matrix	Unitary	K-tuple	4
Mismatch penalty	1	Joining penalty	30
Gap penalty	1.00	Window size	8
Gap size penalty	0.33		
Cutoff score	0		
Randomization group	0		

SEARCH STATISTICS

	Mean	Median	Standard Deviation
Scores:	6	7	0.89

	CPU	Total Elapsed
Times:	00:00:00.00	00:00:00.00

```
Number of residues:                47269
Number of sequences searched:          6
Number of scores above cutoff:         6

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:
```

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
	**** 1 standard deviation above mean ****					
1. hpu06714	TOIG of: hpu06714 check: 486	7801	7	7	1.12	0
2. a12360	TOIG of: a12360 check: 1580	7909	7	7	1.12	0
	**** 0 standard deviation from mean ****					
3. pph31a	TOIG of: pph31a check: 5866	7912	6	6	0.00	0
4. pph11	TOIG of: pph11 check: 3689	7931	6	6	0.00	0
	**** 1 standard deviation below mean ****					
5. af131950	TOIG of: af131950 check: 557	7812	5	5	-1.12	0
6. pph16	TOIG of: pph16 check: 6074	7904	5	5	-1.12	0

```
1. FRAG2' (1-8)
   hpu06714     TOIG of: hpu06714  check: 4862  from: 1  to: 7801

TOIG of: hpu06714  check: 4862  from: 1  to: 7801

LOCUS       HPU06714     7801 bp    DNA                 VRL      04-FEB-1997
DEFINITION  Human papillomavirus HPV-1A (3-3), complete genome.
ACCESSION   U06714
VERSION     U06714.1  GI:458704
KEYWORDS    .
SOURCE      Human papillomavirus.
  ORGANISM  Human papillomavirus
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 7801)
  AUTHORS   Danos,O., Katinka,M. and Yaniv,M.
  TITLE     Human papillomavirus 1A complete DNA sequence: a novel type of
            genome organization among Papovaviridae
  JOURNAL   EMBO J. 1, 231-236 (1982)
  MEDLINE   84182467
REFERENCE   2  (bases 1 to 7801)
  AUTHORS   Meissner,J.
  TITLE     Complete nucleotide sequencing of an HPV-1a variant and
            determination of extant errors in the prototype HPV-1a sequence
  JOURNAL   Virus Genes 9 (2), 189-191 (1995)
  MEDLINE   95250312
REFERENCE   3  (bases 1 to 7801)
  AUTHORS   Meissner,J.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-FEB-1994) John D. Meissner, Duke University Medical
            Center, Microbiology, 277 Carl Building, Research Drive, Durham, NC
            27710 USA
FEATURES             Location/Qualifiers
     source          1. .7801
                     /organism="Human papillomavirus"
                     /strain="HPV-1A (3-3)"
                     /db_xref="taxon:10566"
     variation       142
                     /replace="a"
     conflict        1283
                     /citation=[1]
                     /replace="a"
     conflict        2301
                     /citation=[1]
                     /replace="t"
     variation       2825
                     /replace="t"
     conflict        3884. .3886
                     /citation=[1]
                     /replace="aa"
     conflict        4332
                     /citation=[1]
                     /replace="a"
     conflict        4376. .4382
                     /citation=[1]
                     /replace="gagggaa"
     variation       5794
                     /replace="a"
     variation       6905
```

```
     variation       /replace="g"
                     7186. .7187
                     /note="15 bp deletion"
                     /citation=[1]
     variation       /replace="ctagattgtattgctat"
                     7560
     variation       /replace="c"
                     7618
     conflict        /replace="g"
                     7677. .7678
                     /citation=[1]
     variation       /replace="cc"
                     7787
                     /replace="g"
BASE COUNT      2389 a    1482 c    1664 g    2266 t
ORIGIN

HPU06714  Length: 7801  November 28, 2001 14:10  Type: N  Check: 4862  ..

Initial Score     =      7  Optimized Score  =      7  Significance =  1.12
Residue Identity  =    87%  Matches          =      7  Mismatches   =     1
Gaps              =      0  Conservative Substitutions               =     0


                                                      X       X
                                                      CGAACGTC
                                                      |||||||
   TCCACTGATGAATATGTAACCAGAACCAATCTCTTCTACCATGCAACATCTGAACGTCTACTGCTGGTCGGA
5500       5510       5520       5530       5540       5550     X 5560

   CATCCTTTGTTTGAGATCTCCAGTAATCAAACTGTA
5570       5580       5590       5600


2. FRAG2' (1-8)
   a12360          TOIG of: a12360  check: 1580  from: 1  to: 7909

TOIG of: a12360  check: 1580  from: 1  to: 7909

LOCUS       A12360          7909 bp     DNA                     PAT       12-DEC-1993
DEFINITION  complete nucleotide sequence of HPV-33.
ACCESSION   A12360
VERSION     A12360.1  GI:492936
KEYWORDS    .
SOURCE      Human papillomavirus type 33.
  ORGANISM  Human papillomavirus type 33
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 7909)
  AUTHORS   .
  JOURNAL   Patent: WO 8705630-A 1 24-SEP-1987;
FEATURES             Location/Qualifiers
     source          1. .7909
                     /organism="Human papillomavirus type 33"
                     /db_xref="taxon:10586"
BASE COUNT      2544 a    1354 c    1535 g    2474 t      2 others
ORIGIN

A12360  Length: 7909  November 28, 2001 14:10  Type: N  Check: 1580  ..

Initial Score     =      7  Optimized Score  =      7  Significance =  1.12
Residue Identity  =    87%  Matches          =      7  Mismatches   =     1
Gaps              =      0  Conservative Substitutions               =     0


                                                    X       X
                                                    CGAACGTC
                                                    |||||||
   GTTCACAAAGTGCTGCGGAGGACGTTGTTGATCGTGCTGCAAACCCGTGTAGAACGTCTATTAATAAAAATA
1160       1170       1180       1190       1200     X 1210       1220

   AAGAATGCACATACAGAAAACGAAAAATAGATGAGC
1230       1240       1250       1260


3. FRAG2' (1-8)
   pph31a          TOIG of: pph31a  check: 5866  from: 1  to: 7912

TOIG of: pph31a  check: 5866  from: 1  to: 7912

LOCUS       PPH31A          7912 bp     DNA     circular  VRL       18-MAR-1994
DEFINITION  Human papillomavirus type 31 (HPV-31) complete genome.
ACCESSION   J04353
VERSION     J04353.1  GI:333048
KEYWORDS    complete genome.
SOURCE      Human papillomavirus type 31 DNA.
  ORGANISM  Human papillomavirus type 31
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 7912)
  AUTHORS   Goldsborough,M.D., DiSelvestre,D., Temple,G.F. and Lorincz,A.T.
  TITLE     Nucleotide sequence of human papillomavirus type 31: A cervical
            neoplasia associated virus
  JOURNAL   Virology 171, 306-311 (1989)
  MEDLINE   89299478
COMMENT     Draft entry and computer-readable copy of sequence [1] kindly
            submitted by M.D.Goldsborough, 05-JUL-1989.
FEATURES             Location/Qualifiers
     source          1. .7912
                     /organism="Human papillomavirus type 31"
                     /db_xref="taxon:10585"
     TATA_signal     19. .24
     TATA_signal     69. .74
     gene            108. .557
     CDS             108. .557
                     /gene="E6"
                     /gene="E6"
                     /note="ORF E6 from bp 39 to 557"
                     /codon_start=1
                     /product="transforming protein"
                     /protein_id="AAA46950.1"
                     /db_xref="GI:459916"
                     /translation="MFKNPAERPRKLHELSSALEIPYDELRLNCVYCKGQLTETEVLD
                     FAFTDLTIVYRDDTPHGVCTKCLRFYSKVSEFRWYRYSVYGTTLEKLTNKGICDLLIR
     misc_feature    CITCQRPLCPEEKQRHLDKKKRFHNIGGRWTGRCIACWRRPRTETQV"
                     228. .236
                     /gene="E6"
     misc_feature    /standard_name="Splice donor"
                     403. .414
                     /gene="E6"
     gene            /standard_name="Splice acceptor"
                     560. .856
     CDS             /gene="E7"
                     560. .856
                     /gene="E7"
                     /note="ORF E7 from bp 545 to 856"
                     /codon_start=1
                     /product="transforming protein"
                     /protein_id="AAA46951.1"
                     /db_xref="GI:459917"
                     /translation="MRGETPTLQDYVLDLQPEATDLHCYEQLPDSSDEEDVIDSPAGQ
     gene            AEPDTSNYNIVTFCCQCKSTLRLCVQSTQVDIRILQELLMGSFGIVCPNCSTRL"
                     862. .2751
     CDS             /gene="E1"
                     862. .2751
                     /gene="E1"
                     /note="ORF E1 from bp 850 to bp 2751"
                     /codon_start=1
                     /product="replication protein"
                     /protein_id="AAA46952.1"
                     /db_xref="GI:459918"
                     /translation="MADPAGTDGEGTGCNGWFYVEAVIDRQTGDNISEDENEDSSDTG
                     EDMVDFIDNCNVYNNQAEAETAQALFHAQEAEEHAEAVQVLKRKYVGSPLSDISSCVD
                     YNISPRLKAICIENNSKTAKRRLFELPDSGYGNTEVETQQMVQVEEQQTTLSCNGSDG
                     THSERENETPTRNILQVLKTSNGKAAMLGKFKELYGVSFMELIRPFQSNKSTCTDWCV
                     AAFGVTGTVAEGFKTLLQPYCLYCHLQSLACSWGMVMLMLVRFKCAKNRITIEKLLEK
                     LLCISTNCMLIQPPKLRSTAAALYWYRTGMSNISDVYGETPEWIERQTVLQHSFNDTT
                     FDLSQMVQWAYDNDVMDDSEIAYKYAQLADSDSNACAFLKSNSQAKIVKDCGTMCRHY
```

```
                KRAEKRQMSMGQWIKSRCDKVSDEGDWRDIVKFLRYQQIEFVSFLSALKLFLKGVPKK
                NCILIHGAPNTGKSYFGMSLISFLQGCIISYANSKSHFWLQPLADAKIGMLDDATTPC
                WHYIDNYLRNALDGNPVSIDVKHKALMQLKCPPLLITSNINAGKDDRWPYLHSRLVVF
                TFPNPFPFDKNGNPVYELSDKNWKSFFSRTWCRLNLHEEEDKENDGDSFSTFKCVSGQ
                NIRTL"
gene            2693. .3811
                /gene="E2"
CDS             2693. .3811
                /gene="E2"
                /note="ORF E2 from bp 2663 to 3811"
                /codon_start=1
                /product="regulatory protein"
                /protein_id="AAA46953.1"
                /db_xref="GI:459919"
                /translation="METLSQRLNVCQDKILEHYENDSKRLCDHIDYWKHIRLECVLMY
                KAREMGIHSINHQVVPALSVSKAKALQAIELQMMLETLNNTEYKNEDWTMQQTSLELY
                LTAPTGCLKKHGYTVEVQFDGDVHNTMHYTNWKFIYLCIDGQCTVVEGQVNCKGIYYV
                HEGHITYFVNFTEEAKKYGTGKKWEVHAGGQVIVFPESVFSSDEISFAGIVTKLPTAN
                NTTTSNSKTCALGTSEGVRRATTSTKRPRTEPEHRNTHHPNKLLRGDSVDSVNCGVIS
                AAACTNQTRAVSCPATTPIIHLKGDANILKCLRYRLSKYKQLYEQVSSTWHWTCTDGK
                HKNAIVTLTYISTSQRDDFLNTVKIPNTVSVSTGYMTI"
gene            3270. .3578
                /gene="E4"
CDS             <3270. .3578
                /gene="E4"
                /note="ORF E4 from bp 3270 to 3578"
                /codon_start=1
                /protein_id="AAA46949.1"
                /db_xref="GI:459915"
                /translation="LFFLNLYLAVTKYPLLGLLQSYQQPTTPPHRIPKPAPWAPVKVC
                GGRRRLLSDQEQSQSTETPTTPTSCCEATPWTVSTVGLSVQLHAQTKQGLSVVLQLHL
                "
gene            3816. .4070
                /gene="E5"
CDS             3816. .4070
                /gene="E5"
                /note="ORF E5 from bp 3804 to 4070"
                /codon_start=1
                /protein_id="AAA46954.1"
                /db_xref="GI:459920"
                /translation="MIELNISTVSIVLCFLLCFCVLLFVCLVIRPLVLSVSVYATLLL
                LIVILWVIATSPLRCFCIYVVFIYIPLFVIHTHASFLSQQ"
repeat_region   4099. .4134
polyA_signal    4138. .4143
                /note="putative"
repeat_region   4143. .4158
gene            4171. .5571
                /gene="L2"
CDS             4171. .5571
                /gene="L2"
                /note="ORF L2 from bp 4060 to bp 5571"
                /codon_start=1
                /product="minor capsid protein"
                /protein_id="AAA46955.1"
                /db_xref="GI:459921"
                /translation="MRSKRSTKRTKRASATQLYQTCKAAGTCPSDVIPKIEHTTIADQ
                ILRYGSMGVFFGGLGIGSGSGTGGRTGYVPLSTRPSTVSEASIPIRPPVSIDPVGPLD
                PSIVSLVEESGIVDVGAPAPIPHPPTTSGFDIATTADTTPAILDVTSVSTHENPTFTD
                PSVLQPPTPAETSGHLLLSSSSISTHNYEEIPMDTFIVSTNNENITSSTPIPGVRRPA
                RLGLYSKATQQVKVIDPTFLSAPKQLITYENPAYETVNAEESLYFSNTSHNIAPDPDF
                LDIIALHRPALTSRRNTVRYSRLGNKQTLRTRSGATIGARVHYYYDISSINPAGESIE
                MQPLGASATTTSTLNDGLYDIYADTDFTVDTPATHNVSPSTAVQSTSAVSAYVPTNTT
                VPLSTGFDIPIFSGPDVPIEHAPTQVFPFPLAPTTPQVSIFVDGGDFYLHPSYYMLKR
                RRKRVSYFFTDVSVAA"
gene            5552. .7066
                /gene="L1"
CDS             5552. .7066
                /gene="L1"
                /note="ORF L1 from bp 5516 to  7066"
                /codon_start=1
                /product="major capsid protein"
                /protein_id="AAA46956.1"
                /db_xref="GI 459922"
                /translation="MSLWRPSEATVYLPPVPVSKVVSTDEYVTRTNIYYHAGSARLLT
                VGHPYYSIPKSDNPKKIVVPKVSGLQYRVFRVRLPDPNKFGFPDTSFYNPETQRLVWA
                CVGLEVGRGQPLGVGISGHPLLNKFDDTENSNRYAGGPGTDNRECISMDYKQTQLCLL
                GCKPPIGEHWGKGSPCSNNAITPGDCPPLELKNSVIQDGDMVDTGFGAMDFTALQDTK
                SNVPLDICNSICKYPDYLKMVAEPYGDTLFFYLRREQMFVRHFFNRSGTVGESVPTDL
                YIKGSGSTATLANSTYFPTPSGSMVTSDAQIFNKPYWMQRAQGHNNGICWGNQLFVTV
                VDTTRSTNMSVCAAIANSDTTFKSSNFKEYLRHGEEFDLQFIFQLCKITLSADIMTYI
                HSMNPAILEDWNFGLTTPPSGSLEDTYRFVTSQAITCQKTAPQKPKEDPFKDYVFWEV
                NLKEKFSADLDQFPLGRKFLLQAGYRARPKFKAGKRSAPSASTTTPAKRKKTKK"
polyA_signal    7227. .7231
                /note="putative"
repeat_region   7291. .7302
repeat_region   7314. .7333
                /note="putative"
protein_bind    7406. .7420
                /standard_name="glucocorticoid responsive element"
                /bound_moiety="hormone receptor"
protein_bind    7477. .7488
                /function="gene transcription"
                /bound_moiety="E2"
enhancer        7542. .7549
                /standard_name="keratinocyte-dependant enhancer"
protein_bind    7868. .7879
                /function="gene transcription"
                /bound_moiety="E2"

BASE COUNT      2528 a   1364 c   1572 g   2448 t
ORIGIN

PPH31A  Length: 7912  November 28, 2001 14:10  Type: N  Check: 5866  ..

Initial Score   =     6  Optimized Score  =     6  Significance  =  0.00
Residue Identity =   75%  Matches         =     6  Mismatches    =     2
Gaps            =     0  Conservative Substitutions              =     0

                                                        X      X
                                                        CGAACGTC
                                                          ||||||

   ATTTTTGTTGATGGGGGTGATTTTTATTTGCACCCTAGTTATTATATGTTAAAACGTCGACGTAAACGTGTA
     5470       5480       5490       5500       5510   X  5520       5530

   TCATATTTTTTTACAGATGTCTCTGTGGCGGCCTAG
    5540       5550       5560       5570


4. FRAG2' (1-8)
   pph11        TOIG of: pph11  check: 3689  from: 1  to: 7931

TOIG of: pph11  check: 3689  from: 1  to: 7931

LOCUS       PPH11          7931 bp    DNA   circular  VRL      02-JUN-1994
DEFINITION  Human papillomavirus type 11 (HPV-11) complete genome.
ACCESSION   M14119
VERSION     M14119.1  GI:333026
KEYWORDS    complete genome.
SOURCE      Human laryngeal papillomavirus type 11 DNA.
  ORGANISM  Human papillomavirus type 11
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 7931)
  AUTHORS   Dartmann,K., Schwarz,E., Gissmann,L. and Zur Hausen,H.
  TITLE     The nucleotide sequence and genome organization of human papilloma
            virus type 11
  JOURNAL   Virology 151, 124-130 (1986)
  MEDLINE   86181601
COMMENT     ORF L1 is assumed to encode the major structural protein.
FEATURES             Location/Qualifiers
     source          1. .7931
                     /organism="Human papillomavirus type 11"
                     /db_xref="taxon:10580"
     CAAT_signal     9. .15
                     /note="putative"
     protein_bind    35. .46
                     /note="putative"
```

```
                     /function="gene transcription"
     protein_bind    /bound_moiety="E2"
                     50. .61
                     /note="putative"
                     /function="gene transcription"
     TATA_signal     /bound_moiety="E2"
                     66. .71
     gene            /note="putative"
                     102. .554
     CDS             /gene="E6"
                     102. .554
                     /gene="E6"
                     /note="102 is position of first start codon in ORF E6;
                     putative"
                     /codon_start=1
                     /product="transforming protein"
                     /protein_id="AAA46927.1"
                     /db_xref="GI:496193"
                     /translation="MESKDASTSATSIDQLCKTFNLSLHTLQIQCVFCRNALTTAEIY
                     AYAYKNLKVVWRDNFPFAACACCLELQGKINQYRHFNYAAYAPTVEEETNEDILKVLI
     gene            RCYLCHKPLCEIEKLKHILGKARFIKLNNQWKGRCLHCWTTCMEDLLP"
                     530. .826
     CDS             /gene="E7"
                     530. .826
                     /gene="E7"
                     /note="530 is position of first start codon in ORF E7;
                     putative"
                     /codon_start=1
                     /product="transforming protein"
                     /protein_id="AAA46928.1"
                     /db_xref="GI:496194"
                     /translation="MHGRLVTLKDIVLDLQPPDPVGLHCYEQLEDSSEDEVDKVDKQD
     gene            AQPLTQHYQILTCCCGCDSNVRLVVECTDGDIRQLQDLLLGTLNIVCPICAPKP"
                     832. .2781
     CDS             /gene="E1"
                     832. .2781
                     /gene="E1"
                     /note="832 is position of first start codon in ORF E1;
                     putative"
                     /codon_start=1
                     /product="replication protein"
                     /protein_id="AAA46929.1"
                     /db_xref="GI:496195"
                     /translation="MADDSGTENEGSGCTGWFMVEAIVEHTTGTQISEDEEEEVEDSG
                     YDMVDFIDDRHITQNSVEAQALFNRQEADAHYATVQDLKRKYLGSPYVSPISNVANAV
                     ESEISPRLDAIKLTTQPKKVKRRLFETRELTDSGYGYSEVEAATQVEKHGDPENGGDG
                     QERDTGRDIEGEGVEHREAEAVDDSTREHADTSGILELLKCKDIRSTLHGKFKDCFGL
                     SFVDLIRPFKSDRTTCADWVVAGFGIHHSIADAFQKLIEPLSLYAHIQWLTNAWGMVL
                     LVLIRFKVNKSRCTVARTLGTLLNIPENHMLIEPPKIQSGVRALYWFRTGISNASTVI
                     GEAPEWITRQTVIEHSLADSQFKLTEMVQWAYDNDICEESEIAFEYAQRGDFDSNARA
                     FLNSNMQAKYVKDCAIMCRHYKHAAEMKKMSIKQWIKYRGTKVDSVGNWKPIVQFLRHQ
                     NIEFIPFLSKLKLWLHGTPKKNCIAIVGPPDTGKSCFCMSLIKFLGGTVISYVNSCSH
                     FWLQPLTDAKVALLDDATQPCWTYMDTYMRNLLDGNPMSIDRKHRALTLIKCPPLLVT
     gene            SNIDISKEEKYKYLHSRVTTFTFPNPFPFDRNGNAVYELSDANWKCFFERLSSSLDIE
                     DSEDEEDGSNSQAFRCVPGSVVRTL"
     CDS             2723. .3826
                     /gene="E2"
                     2723. .3826
                     /gene="E2"
                     /note="2723 is position of first start codon in ORF E2;
                     putative"
                     /codon_start=1
                     /product="regulatory protein"
                     /protein_id="AAA46930.1"
                     /db_xref="GI:496196"
                     /translation="MEAIAKRLDACQDQLLELYEENSIDIHKHIMHWKCIRLESVLLH
                     KAKQMGLSHIGLQVVPPLTVSETKGHNAIEMQMHLESLAKTQYGVEPWTLQDTSYEMW
                     LTPPKRCFKKQGNTVEVKFDGCEDNVMEYVVWTHIYLQDNDSWVKVTSSVDAKGIYYT
                     CGQFKTYYVNFNKEAQKYGSTNHWEVCYGSTVICSPASVSSTVREVSIAEPTTYTPAQ
                     TTAPTVSACTTEDGVSAPPRKRARGPSTNNTLCVANIRSVDSTINNIVTDNYNKHQRR
     gene            NNCHSAATPIVQLQGDSNCLKCFRYRLNDKYKHLFELASSTWHWASPEAPHKNAIVTL
                     TYSSEEQRQQFLNSVKIPPTIRHKVGFMSLHLL"
     CDS             3255. .3581
```

```
     CDS             /gene="E4"
                     3255. .3581
                     /gene="E4"
                     /note="3255 is position of first start codon in ORF E4;
                     putative"
                     /codon_start=1
                     /protein_id="AAA46931.1"
                     /db_xref="GI:496197"
                     /translation="MVVPIIGKYVMAAQLYVLLHLYLALYEKYPLLNLLHTPPHRPPP
                     LQCPPAPRKTACRRRLGSEHVDRPLTTPCVWPTSDPWTVQSTTSSLTITTSTKEGTTV
     gene            TVQLRL"
                     3871. .4146
     CDS             /gene="E5A"
                     3871. .4146
                     /gene="E5A"
                     /note="3871 is position of first start codon in ORFE5A;
                     putative"
                     /codon_start=1
                     /protein_id="AAA46932.1"
                     /db_xref="GI:496198"
                     /translation="MEVVPVQIAAATTTTLILPVVIAFAVCILSIVLIILISDFVVYT
     gene            SVLVLTLLLYLLLWLLLTTPLQFFLLTLCVCYFPAFYIHIYIVQTQQ"
                     4146. .4370
     CDS             /gene="E5B"
                     4146. .4370
                     /gene="E5B"
                     /note="4146 is position of first start codon in ORF E5B;
                     putative"
                     /codon_start=1
                     /protein_id="AAA46933.1"
                     /db_xref="GI:496199"
                     /translation="MVMLTCHLNDGDTWLFLWLFTAFVVAVLGLLLLHYRAVHGTEKT
     polyA_signal    KCAKCKSNRNTTVDYVYMSHGDNGDYVYMN"
                     4371. .4376
     gene            /note="putative"
                     4417. .5784
     CDS             /gene="L2"
                     4417. .5784
                     /gene="L2"
                     /note="4417 is position of first start codon in ORF L2;
                     putative"
                     /codon_start=1
                     /product="minor capsid protein"
                     /protein_id="AAA46934.1"
                     /db_xref="GI:496200"
                     /translation="MKPRARRRKRASATQLYQTCKATGTCPPDVIPKVEHTTIADQIL
                     KWGSLGVFFGGLGIGTGAGSGGRAGYIPLGSSPKPAITGGPAARPPVLVEPVAPSDPS
                     IVSLIEESAIINAGAPEVVPPTQGGFTITSSESTTPAILDVSVTNHTTTSVFQNPLFT
                     EPSVIQPQPPVEASGHILISAPTITSQHVEDIPLDTFVVSSSDSGPTSSTPLPRAFPR
                     PRVGLYSRALQQVQVTDPAFLSTPQRLVTYDNPVYEGEDVSLQFTHESIHNAPDEAFM
                     DIIRLHRPAITSRRGLVRFSRIGQRGSMYTRSGQHIGARIHYFQDISPVTQAAEEIEL
                     HPLVAAENDTFDIYAEPFDPIPDPVQHSVTQSYLTSTPNTLSQSWGNTTVPLSIPSDW
                     FVQSGPDITFPTASMGTPFSPVTPALPTGPVFITGSDFYLHPTWYFARRRRKRIPLFF
     polyA_signal    TDVAA"
                     4545. .4551
                     /gene="L2"
     gene            /note="putative"
                     5771. .7276
     CDS             /gene="L1"
                     5771. .7276
                     /gene="L1"
                     /note="5771 is position of first start codon in ORF L1;
                     putative"
                     /codon_start=1
                     /product="major capsid protein"
                     /protein_id="AAA46935.1"
                     /db_xref="GI:496201"
                     /translation="MWRPSDSTVYVPPPNPVSKVVATDAYVKRTNIFYHASSSRLLAV
                     GHPYYSIKKVNKTVVPKVSGYQYRVFKVVLPDPNKFALPDSSLFDPTTQRLVWACTGL
                     EVGRGQPLGVGVSGHPLLNKYDDVENSGGYGGNPGQDNRVNVGMDYKQTQLCMVGCAP
                     PLGEHWGKGTQCSNTSVQNGDCPPLELITSVIQDGDMVDTGFGAMNFADLQTNKSDVP
                     LDICGTVCKYPDYLQMAADPYGDRLFFYLRKEQMFARHFFNRAGTVGEPVPDDLLVKG
                     GNNRSSVASSIYVHTPSGSLVSSEAQLFNKPYWLQKAQGHNNGICWGNHLFVTVVDTT
```

```
                RSTNMTLCASVSKSATYTNSDYKEYMRHVEEFDLQFIFQLCSITLSAEVMAYIHTMNP
                SVLEDWNFGLSPPPNGTLEDTYRYVQSQAITCQKPTPEKEKQDPYKDMSFWEVNLKEK
                FSSELDQFPLGRKFLLQSGYRGRTSARTGIKRPAVSKPSTAPKRKRTKTKK"
5'UTR           join(7277. .7931,1. .101)
                /standard_name="LCR"
                /function="regulatory region"
repeat_region   7320. .7360
                /standard_name="direct repeats"
                /note="putative"
repeat_region   7339. .7374
                /standard_name="direct repeats"
                /note="putative"
repeat_region   7374. .7403
                /note="putative"
polyA_signal    7457. .7462
                /note="putative"
protein_bind    7592. .7603
                /note="putative"
                /function="gene transcription"
                /bound_moiety="E2"
polyA_signal    7748. .7753
                /note="putative"
protein_bind    7890. .7901
                /note="putative"
                /function="gene transcription"
                /bound_moiety="E2"
BASE COUNT      2406 a   1519 c   1736 g   2270 t
ORIGIN      4557 bp upstream of HindIII site.

  PPH11  Length: 7931  November 28, 2001 14:10  Type: N  Check: 3689  ..

Initial Score  =      6  Optimized Score  =      6  Significance  =   0.00
Residue Identity =   75%  Matches         =      6  Mismatches    =      2
Gaps           =      0  Conservative Substitutions               =      0

                                                      X       X
                                                      CGAACGTC
                                                        ||||||
     ATGCACCTGATGAAGCATTTATGGATATTATTAGACTACATAGACCAGCTATAACGTCCAGACGGGGTCTTG
        5230       5240       5250       5260       5270 X     5280       5290

     TGCGTTTTAGTCGCATTGGGCAACGGGGGTCCATGT
        5300       5310       5320


5. FRAG2' (1-8)
af131950      TOIG of: af131950  check: 557  from: 1  to: 7812

  TOIG of: af131950  check: 557  from: 1  to: 7812

LOCUS       AF131950     7812 bp    DNA             VRL       07-FEB-2001
DEFINITION  Human papillomavirus candHPV85, complete genome.
ACCESSION   AF131950
VERSION     AF131950.1  GI:4574720
KEYWORDS    .
SOURCE      Human papillomavirus candHPV85
  ORGANISM  Human papillomavirus candHPV85
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 7812)
  AUTHORS   Chow,V.T.K. and Leong,P.W.F.
  TITLE     Complete nucleotide sequence, genomic organization and phylogenetic
            analysis of a novel genital human papillomavirus type, HLT7474-S
  JOURNAL   J. Gen. Virol. 80 (Pt 11), 2923-2929 (1999)
  MEDLINE   20047972
   PUBMED   10580054
REFERENCE   2  (bases 1 to 7812)
  AUTHORS   Chow,V.T.K. and Leong,W.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-FEB-1999) Department of Microbiology, National
            University of Singapore, 10 Kent Ridge Crescent, Singapore 119260,
            Singapore
FEATURES             Location/Qualifiers
     source          1. .7812
                     /organism="Human papillomavirus candHPV85"
                     /db_xref="taxon:151757"
                     /note="isolated from scraped uterine cervical cells from a
                     female sex worker; overlapping PCR products"
     protein_bind    40. .51
                     /note="putative"
                     /bound_moiety="E2 protein"
                     /function="transcriptional regulation"
     protein_bind    56. .67
                     /note="putative"
                     /bound_moiety="E2 protein"
                     /function="transcriptional regulation"
     TATA_signal     71. .79
                     /note="putative"
     gene            105. .578
                     /gene="E6"
     CDS             105. .578
                     /gene="E6"
                     /codon_start=1
                     /product="putative transforming protein E6"
                     /protein_id="AAD24181.1"
                     /db_xref="GI:4574721"
                     /translation="MAEFGNPATRPYKLPDLCNTLDTSLQDIEISCVYCKSVLQRTEV
                     YEFAFADLFVVYRDGIPYAACQNCLMFYSKIRELRYYSDSVYGETLEKLTNSNIYDIL
                     IRCLRCQKPLCPAEKLKHLNEKRRFHKIAGKYRGQCRRCMTRAQEQQGSRRETQV"
     gene            587. .913
                     /gene="E7"
     CDS             587. .913
                     /gene="E7"
                     /codon_start=1
                     /product="putative transforming protein E7"
                     /protein_id="AAD24182.1"
                     /db_xref="GI:4574722"
                     /translation="MHGPKPTVHDIVLDLEPYNEVQEVDLYCYEELNNSEEEIDEPDN
                     AINHRQPLLARREELQRHTICCVCCKCEASLQLVVESSAADLRDLQQLFLGTLSFLCP
                     LCAVLR"
     gene            920. .2866
                     /gene="E1"
     CDS             920. .2866
                     /gene="E1"
                     /codon_start=1
                     /product="putative replication protein E1"
                     /protein_id="AAD24183.1"
                     /db_xref="GI:4574723"
                     /translation="MADPEGTDGDGTGCNGWFFVQAIVDKKTGDTISEDEDEDGTDTG
                     SDLVDFIDTSNTNYMQADREAAQALLHAQEVETDTKLLHALKRKYGAHSTENSPCRDT
                     ASIHSNLSSPLQEISLNSYNNTAKRRLCSVPDSGYGNTQVETLQTQVTLDTNVFGDGK
                     NGDGLNSEACSTDNEMDIENQNPNSPMTQIVSLLKVNNKKAAILAKFKETYGLSFTDL
                     VRTFKSDKTTCTDWVAAICGVNPNIAEGFKTLIQPYVLYAHIQCMDCSWGVFILALLR
                     YKCGKNRLTVAKGLSTLLHVPDTHMLIEPPKLRSSCAALYWYRTGISNISEVTGDTPE
                     WIQRQTIIQHGIDDSVFDLSEMIQWAFDNDYIDESDIAYEYAQLADCNSNAAAFLKSN
                     CQAKYLRDCAVMCRHYKRAQRKQMNMSQWISYRCDKIDDGGDWKPIVQFLRFQGIEFI
                     TFLRAFKDFLKGTPKKNCIVIYGPANTGKSYFCMSLIQFLHGTVLSFVNSNSHFWLEP
                     LTDTKIAMVDDATPTCWSYFDNYMRNALDGNPISIDRKHKHLIQMKCPPMLITSNTNP
                     ATDDRWPYLRSRVTVFTFPHTFPFDSNGNPVYDINDKNWKCFFKRTWSRLDLHQEEED
                     TENDGCPIPTFKCVTGKDSRTLRD"
     gene            2799. .3932
                     /gene="E2"
     CDS             2799. .3932
                     /gene="E2"
                     /codon_start=1
                     /product="putative regulatory protein E2"
                     /protein_id="AAD24184.1"
                     /db_xref="GI:4574724"
                     /translation="MMDALSQRLNVLQEKILEHYETDSTDICDQIDYWKCVRLENAIY
                     YAAREHGLKTINHHVVPTFQISKSKAHEAIELQMALESLAKSEFKNELWTLQDTCQEL
                     YQTPPQQCFKKQGQTVEVRYDGDKDNTMHYTSWDYIYYVTEGDKWCKTKGYVNYCGLY
                     YIKEGQQTYYVQFKCDAQQYGQSGKWEVWYNGKKIECVNTCDKIIQCSESMYSTCDET
                     VSATAIARELQHPTTPYTEATTVCTQKSGGSAPTRNPFRHCGFTETSEVDGLSVDHLN
                     NPLLSAPAGNNGFRKHNSGDTTPIVHLKGDKNRLKCLRYRLQKFHNLYTQISCTWHWI
                     GGKETTRTGILTLAYANETQRQKFLDVVKIPNTVNVSLGYMTL"
     gene            <3412. .3699
```

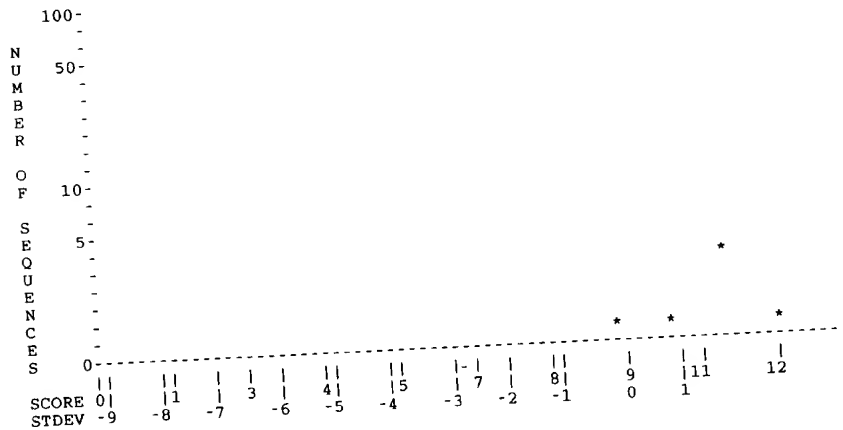
```
     CDS             /gene="E4"
                     <3412. .3699
                     /gene="E4"
                     /note="start codon absent"
                     /codon_start=1
                     /product="putative protein E4"
                     /protein_id="AAD24185.1"
                     /db_xref="GI:4574725"
                     /translation="YNVLNLCTVPVTRQYPLLQLLENYNTPPHRIPKPPPCAPKRAGV
     gene            RRRLETHSDIVDSQKQVKSTDCPWTTSTTPCLVHLQATTGSGSTIVVTLRL"
                     3940. .4161
     CDS             /gene="E5"
                     3940. .4161
                     /gene="E5"
                     /codon_start=1
                     /product="putative protein E5"
                     /protein_id="AAD24186.1"
                     /db_xref="GI:4574726"
                     /translation="MLIAVFLLCFCVCLYVCCTIPLFPATNVCAHAWLLVFLLIVFIT
     gene            TPFTAFIVYIFFFILPMFLLHLHSVHTFD"
                     4226. .5629
     CDS             /gene="L2"
                     4226. .5629
                     /gene="L2"
                     /codon_start=1
                     /product="putative minor capsid protein L2"
                     /protein_id="AAD24187.1"
                     /db_xref="GI:4574727"
                     /translation="MVSHRAARRKRASATDLYKTCKQSGTCPPDVINKVEGTTLADKL
                     LQWTSLGIFLGGLGIGTGSGTGGRTGYIPLGGRPNTVVDVSPARPPVVIESVGPSDPS
                     IVTLVEESSIVTSGAPVPTFTGTSGFEITSSATTTPAVLDITPASGSVQLSSTSFTNP
                     AFTDPSVIEVPQTGEVSGDIFITTTPTSGTHGYEEIPMHTFATQGRGTEPISSTPIPGV
                     RRVAGPRLYSQAYQQVKITNSDFISRPSTLVTFTNPAYEPIDTTLTFSPQDVVPDPDF
                     MDIVRLHRPALTSRRGTVRFSRLGKKLTMSTRSGKQIGAQVHYYHDISPISHIGESIE
                     MQPLLPDAAVTADTNGLFDIYADTDIDNNAMLYDRNISDVTQPTTSTISSVSSRYSNT
                     TIPLATSWDVPVHTGPDMTLPTTIPQWPNIVPLLPNNTHSVVLQGTNYYLWPNYYFIF
     gene            KKRKRVPYFLTDGFVAF"
                     5610. .7124
     CDS             /gene="L1"
                     5610. .7124
                     /gene="L1"
                     /codon_start=1
                     /product="putative major capsid protein L1"
                     /protein_id="AAD24188.1"
                     /db_xref="GI:4574728"
                     /translation="MALWRSSDSKVYLPPPSVAKVVNTDDYVTRTSTFYHAGSSRLLT
                     VGHPYYKVTSNGGRKQDIPKVSAYQYRVFRVTLPDPNKFGLPDTNIYNPETQRLVWAC
                     VGMEVGRGQPLGVSLSGHPFYNKLDDTENSHVATSVVTHDTRDNVSVDYKQTQLCIIG
                     CVPAIGEHWAKGTACKPGAVQTGDCPPLELVNTPIEDGDMIDTGYGAMDFSTLQDNKS
                     EVPLDICQSICKYPDYLQMSADAYGDSMFFCLRREQLFARHFWNRGGTIGDAVPETLY
                     IKGTNDRATPGSCIYSPSPSGSMVSSDAQMFNKPYWLHKAQGQNNGICWHNQLFITVV
                     DTTRSTNLTLSTATTNPVPSIYEPSKFKEYTRHVEEYDLQFIFQLCKITLTTDVMSYI
                     HNMDPTILDSWNFGVSPPPSASLVDTYRFLQSSAITCQKDVVVPQKKDPYEKLKFWNV
     protein_bind    DLKEHFSSDLDQFPLGRKFLLQAGLRPKPTIGPRKRVASTSTATRPSKRKRTAK"
                     7456. .7467
                     /note="putative"
                     /bound_moiety="E2 protein"
     misc_feature    /function="transcriptional regulation"
                     7647. .7666
     protein_bind    /note="consensus motif of oncogenic anogenital HPV types"
                     7776. .7787
                     /note="putative"
                     /bound_moiety="E2 protein"
                     /function="transcriptional regulation"
BASE COUNT      2502 a    1391 c    1558 g    2361 t
ORIGIN

AF131950  Length: 7812  November 28, 2001 14:10  Type: N  Check: 557  ..

Initial Score  =      5  Optimized Score  =      5  Significance = -1.12
Residue Identity =   58%  Matches         =      7  Mismatches   =      1
Gaps           =      4  Conservative Substitutions              =      0
```

```
                                                                    X          X
                                                                    CGAAC----GTC
     CACTGCAAGATATAGAAATAAGCTGTGTATATTGTAAAAGCGTCCTGCAGCGAACAGAGGTATATGAATTTG
        180       190       200       210       220       X  230       X 240
     CCTTTGCTGATTTATTTGTAGTATATAGAGACGGTATACC
      250       260       270       280

  6. FRAG2' (1-8)
     pph16          TOIG of: pph16  check: 6074  from: 1  to: 7904

  TOIG of: pph16  check: 6074  from: 1  to: 7904

LOCUS       PPH16           7904 bp    DNA   circular  VRL       18-MAR-1994
DEFINITION  Human papillomavirus type 16 (HPV16), complete genome.
ACCESSION   K02718
VERSION     K02718.1  GI:333031
KEYWORDS    circular; complete genome.
SOURCE      Papilloma virus type 16 DNA, isolated from a human invasive
              cervical carcinoma.
  ORGANISM  Human papillomavirus type 16
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 7904)
  AUTHORS   Seedorf,K., Kraemmer,G., Duerst,M., Suhai,S. and Roewekamp,W.G.
  TITLE     Human papillomavirus type 16 DNA sequence
  JOURNAL   Virology 145, 181-185 (1985)
  MEDLINE   85246220
REFERENCE   2  (sites)
  AUTHORS   Kennedy,I.M., Haddow,J.K. and Clements,J.B.
  TITLE     A negative element in the human poapillomavirus type 16 genome acts
              at the level of late mRNA stability
  JOURNAL   J. Virol. 65, 2093-2097 (1991)
  MEDLINE   91162763
COMMENT     The sense strand of this double-stranded circular genome is shown,
            with a numbering system matching the first 60 bp of HPVal, HPV6b
            and BPV1. The annotation of sites and features is solely based upon
            homology comparison with these other papillomaviruses. In addition
            to the coding sequences reported below, the authors note open
            reading frames which do not start with 'ATG', but which are found
            in other papillomaviruses. In particular, a second portion of the
            E1 gene may be located out to base 2813 (the E1 protein is thought
            to be generally involved in DNA replication).
            A potential 'CAT'-box region is found beginning at base 7895 below,
            and 'TATA' boxes for early and late transcripts may be located at
            17, 65 and 4289. Potential polyadenylation signals are at bases
            4213 and 7260.
            HPV16, in comparison to HPV types 6 and 11, is more often
            associated with malignant genital cancers in humans.
FEATURES              Location/Qualifiers
     source           1. .7904
                      /organism="Human papillomavirus type 16"
                      /db_xref="taxon:10581"
     TATA_signal      17. .23
     TATA_signal      65. .71
     gene             83. .559
     CDS              /gene="E6"
                      83. .559
                      /gene="E6"
                      /note="E6 ORF from 65 to 559; putative"
                      /codon_start=1
                      /product="transforming protein"
                      /protein_id="AAA46939.1"
                      /db_xref="GI:333032"
                      /translation="MHQKRTAMFQDPQERPRKLPQLCTELQTTIHDIILECVYCKQQL
                      LRREVYDFAFRDLCIVYRDGNPYAVCDKCLKFYSKISEYRHYCYSLYGTTLEQQYNKP
     gene             LCDLLIRCINCQKPLCPEEKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL"
                      562. .858
     CDS              /gene="E7"
                      562. .858
                      /gene="E7"
```

```
                /note="E7 ORF from 544 to 858; putative"
                /codon_start=1
                /product="transforming protein"
                /protein_id="AAA46940.1"
                /db_xref="GI:333033"
                /translation="MHGDTPTLHEYMLDLQPETTDLYCYEQLNDSSEEEDEIDGPAGQ
                AEPDRAHYNIVTFCCKCDSTLRLCVQSTHVDIRTLEDLLMGTLGIVCPICSQKP"
gene            join(865. .1140,1140. .2813)
                /gene="E1"
CDS             join(865. .1140,1140. .2813)
                /gene="E1"
                /note="E1 interrupted ORF from 859 to 2813; putative"
                /codon_start=1
                /product="replication protein"
                /protein_id="AAA46936.1"
                /db_xref="GI:459912"
                /translation="MADPAGTNGEEGTGCNGWFYVEAVVEKKTGDAISDDENENDSDT
                GEDLVDFIVNDNDYLTQAETETAHALFTAQEAKQHRDAVQVLKRKYLVSPLSDISGCV
                DNNISPRLKAICIEKQSRAAKRRLFESEDSGYGNTEVETQQMLQVEGRHETETPCSQY
                SGGSGGGCSQYSSGSGGEGVSERHTICQTPLTNILNVLKTSNAKAAMLAKFKELYGVS
                FSELVRPFKSNKSTCCDWCIAAFGLTPSIADSIKTLLQQYCLYLHIQSLACSWGMVVL
                LLVRYKCGKNRETIEKLLSKLLCVSPMCMMIEPPKLRSTAAALYWYKTGISNISEVYG
                DTPEWIQRQTVLQHSFNDCTFELSQMVQWAYDNDIVDDSEIAYKYAQLADTNSNASAF
                LKSNSQAKIVKDCATMCRHYKRAEKKQMSMSQWIKYRCDRVDDGGDWKQIVMFLRYQG
                VEFMSFLTALKRFLQGIPKKNCILLYGAANTGKSLFGMSLMKFLQGSVICFVNSKSHF
                WLQPLADAKIGMLDDATVPCWNYIDDNLRNALDGNLVSMDVKHRPLVQLKCPPLLITS
                NINAGTDSRWPYLHNRLVVFTFPNEFPFDENGNPVYELNDKNWKSFFSRTWSRLSLHE
                DEDKENDGDSLPTFKCVSGQNTNTL"
gene            2755. .3852
                /gene="E2"
CDS             2755. .3852
                /gene="E2"
                /note="E2 ORF from 2725 to 3852; putative"
                /codon_start=1
                /product="regulatory protein"
                /protein_id="AAA46941.1"
                /db_xref="GI:333035"
                /translation="METLCQRLNVCQDKILTHYENDSTDLRDHIDYWKHMRLECAIYY
                KAREMGFKHINHQVVPTLAVSKNKALQAIELQLTLETIYNSQYSNEKWTLQDVSLEVY
                LTAPTGCIKKHGYTVEVQFDGDICNTMHYTNWTHIYICEEASVTVVEGQVDYYGLYYV
                HEGIRTYFVQFKDDAEKYSKNKVWEVHAGGQVILCPTSVFSSNEVSSPEIIRQHLANH
                PAATHTKAVALGTEETQTTIQRPRSEPDTGNPCHTTKLLHRDSVDSAPILTAFNSSHK
                GRINCNSNTTPIVHLKGDANTLKCLRYRFKKHCTLYTAVSSTWHWTGHNVKHKSAIVT
                LTYDSEWQRDQFLSQVKIPKTITVSTGFMSI"
gene            3332. .3619
                /gene="E4"
CDS             <3332. .3619
                /gene="E4"
                /note="E4 ORF from 3332 to 3619; putative"
                /codon_start=1
                /protein_id="AAA46937.1"
                /db_xref="GI:459913"
                /translation="YYVLHLCLAATKYPLLKLLGSTWPTTPPRPIPKPSPWAPKKHRR
                LSSDQDQSQTPETPATPLSCCTETQWTVLQSSLHLTAHTKDGLTVIVTLHP"
gene            3863. .4099
                /gene="E5"
CDS             <3863. .4099
                /gene="E5"
                /note="E5 ORF from 3863 to 4099; putative"
                /codon_start=1
                /protein_id="AAA46938.1"
                /db_xref="GI:459914"
                /translation="YCIHNITGVLFALLCVLLCVCLLIRPLLLSVSTYTSLIILVLLL
                WITAASAFRCFIVYIIFVYIPLFLIHTHARFLIT"
polyA_signal    4213. .4218
                /note="putative"
gene            4235. .5656
                /gene="L2"
CDS             4235. .5656
                /gene="L2"
                /note="L2 ORF from 4133 to 5656; putative"
                /codon_start=1
                /product="minor capsid protein"
                /protein_id="AAA46942.1"
                /db_xref="GI:333036"
                /translation="MRHKRSAKRTKRASATQLYKTCKQAGTCPPDIIPKVEGKTIAEQ
                ILQYGSMGVFFGGLGIGTGSGTGGRTGYIPLGTRPPTATDTLAPVRPPLTVDPVGPSD
                PSIVSLVEETSFIDAGAPTSVPSIPPDVSGFSITTSTDTTPAILDINNTVTTVTTHNN
                PTFTDPSVLQPPTPAETGGHFTLSSSTISTHNYEEIPMDTFIVSTNPNTVTSSTPIPG
                SRPVARLGLYSRTTQQVKVVDPAFVTTPTKLITYDNPAYEGIDVDNTLYFSSNDNSIN
                IAPDPDFLDIVALHRPALTSRRTGIRYSRIGNKQTLRTRSGKSIGAKVHYYYDLSTID
                PAEEIELQTITPSTYTTTSHAASPTSINNGLYDIYADDFITDTSTTPVPSVPSTSLSG
                YIPANTTIPFGGAYNIPLVSGPDIPINITDQAPSLIPIVPGSPQYTIIADAGDFYLHP
                SYYMLRKRRKRLPYFFSDVSLAA"
TATA_signal     4289. .4295
                /gene="L2"
gene            5559. .7154
                /gene="L1"
CDS             5559. .7154
                /gene="L1"
                /note="L1 ORF from 5526 to 7154; putative"
                /codon_start=1
                /product="major capsid protein"
                /protein_id="AAA46943.1"
                /db_xref="GI:333037"
                /translation="MQVTFIYILVITCYENDVNVYHIFFQMSLWLPSEATVYLPPVPV
                SKVVSTDEYVARTNIYYHAGTSRLLAVGHPYFPIKKPNNNKILVPKVSGLQYRVFRIH
                LPDPNKFGFPDTSFYNPDTQRLVWACVGVEVGRGQPLGVGISGHPLLNKLDDTENASA
                YAANAGVDNRECISMDYKQTQLCLIGCKPPIGEHWGKGSPCTNVAVNPGDCPPLELIN
                TVIQDGDMVHTGFGAMDFTTLQANKSEVPLDICTSICKYPDYIKMVSEPYGDSLFFYL
                RREQMFVRHLFNRAGTVGENVPDDLYIKGSGSTANLASSNYFPTPSGSMVTSDAQIFN
                KPYWLQRAQGHNNGICWGNQLFVTVVDTTRSTNMSLCAAISTSETTYKNTNFKEYLRH
                GEEYDLQFIFQLCKITLTADVMTYIHSMNSTILEDWNFGLQPPPGGTLEDTYRFVTQA
                IACQKHTPPAPKEDDPLKKYTFWEVNLKEKFSADLDQFPLGRKFLLQAGLKAKPKFTL
                GKRKATPTTSSTSTTAKRKKRKL"
polyA_signal    7260. .7265
BASE COUNT      2601 a    1377 c    1509 g    2417 t
ORIGIN      Unreported.

PPH16  Length: 7904  November 28, 2001 14:10  Type: N  Check: 6074  ..

Initial Score   =      5  Optimized Score  =      5  Significance  = -1.12
Residue Identity =    62%  Matches          =      5  Mismatches    =      3
Gaps            =      0  Conservative Substitutions                =      0

                                                        X     X
                                                        CGAACGTC
                                                        |||||.

    TGTAGAGGCTGTAGTGGAAAAAAAAACAGGGGATGCTATATCAGATGACGAGAACGAAAATGACAGTGATAC
      930       940       950       960       970   X   980       990

    AGGTGAAGATTTGGTAGATTTTATAGTAAATGATAA
     1000      1010      1020      1030
```

Mon Dec 3 08:02:39 2001

```
> O <
O| |O  IntelliGenetics
> O <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file seq1-inv.res made by sdavid on Wed 28 Nov 101 14:21:55-PST.

Query sequence being compared:          SEQ1' (1-22)
Number of sequences searched:                     6
Number of scores above cutoff:                    6

      Results of the initial comparison of SEQ1' (1-22) with:
        File : hpvcomplete.seq
```



PARAMETERS

Similarity matrix	Unitary	K-tuple	4
Mismatch penalty	1	Joining penalty	30
Gap penalty	1.00	Window size	8
Gap size penalty	0.33		
Cutoff score	1		
Randomization group	0		

SEARCH STATISTICS

	Mean	Median	Standard Deviation
Scores:	10	12	1.03

	CPU	Total Elapsed
Times:	00:00:00.00	00:00:00.00

```
Number of residues:                           47269
Number of sequences searched:                     6
Number of scores above cutoff:                    6

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:
```

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
	**** 1 standard deviation above mean ****					
1. pph11	TOIG of: pph11 check: 3689	7931	12	13	1.94	0
	**** 0 standard deviation from mean ****					
2. af131950	TOIG of: af131950 check: 557	7812	11	13	0.97	0
3. pph16	TOIG of: pph16 check: 6074	7904	11	11	0.97	0
4. pph31a	TOIG of: pph31a check: 5866	7912	11	14	0.97	0
5. a12360	TOIG of: a12360 check: 1580	7909	10	11	0.00	0
6. hpu06714	TOIG of: hpu06714 check: 486	7801	9	12	-0.97	0

1. SEQ1' (1-22)
   pph11 TOIG of: pph11 check: 3689 from: 1 to: 7931

TOIG of: pph11 check: 3689 from: 1 to: 7931

```
LOCUS       PPH11          7931 bp    DNA   circular VRL      02-JUN-1994
DEFINITION  Human papillomavirus type 11 (HPV-11) complete genome.
ACCESSION   M14119
VERSION     M14119.1  GI:333026
KEYWORDS    complete genome.
SOURCE      Human laryngeal papillomavirus type 11 DNA.
  ORGANISM  Human papillomavirus type 11
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 7931)
  AUTHORS   Dartmann,K., Schwarz,E., Gissmann,L. and Zur Hausen,H.
  TITLE     The nucleotide sequence and genome organization of human papilloma
            virus type 11
  JOURNAL   Virology 151, 124-130 (1986)
  MEDLINE   86181601
COMMENT     ORF L1 is assumed to encode the major structural protein.
FEATURES             Location/Qualifiers
     source          1. .7931
                     /organism="Human papillomavirus type 11"
                     /db_xref="taxon:10580"
     CAAT_signal     9. .15
                     /note="putative"
     protein_bind    35. .46
                     /note="putative"
                     /function="gene transcription"
                     /bound_moiety="E2"
     protein_bind    50. .61
                     /note="putative"
                     /function="gene transcription"
                     /bound_moiety="E2"
     TATA_signal     66. .71
                     /note="putative"
     gene            102. .554
                     /gene="E6"
     CDS             102. .554
                     /gene="E6"
                     /note="102 is position of first start codon in ORF E6;
                     putative"
                     /codon_start=1
                     /product="transforming protein"
                     /protein_id="AAA46927.1"
                     /db_xref="GI:496193"
                     /translation="MESKDASTSATSIDQLCKTFNLSLHTLQIQCVFCRNALTTAEIY
                     AYAYKNLKVVWRDNFPFAACACCLELQGKINQYRHFNYAAYAPTVEEETNEDILKVLI
                     RCYLCHKPLCE1EKLKHILGKARFIKLNNQWKGRCLHCWTTCMEDLLP"
     gene            530. .826
                     /gene="E7"
     CDS             530. .826
                     /gene="E7"
                     /note="530 is position of first start codon in ORF E7;
                     putative"
                     /codon_start=1
                     /product="transforming protein"
                     /protein_id="AAA46928.1"
                     /db_xref="GI:496194"
```

```
                     /translation="MHGRLVTLKDIVLDLQPPDPVGLHCYEQLEDSSEDEVDKVDKQD
     gene            AQPLTQHYQILTCCCGCDSNVRLVVECTDGDIRQLQDLLLGTLNIVCPICAPKP"
                     832. .2781
     CDS             /gene="E1"
                     832. .2781
                     /gene="E1"
                     /note="832 is position of first start codon in ORF E1;
                     putative"
                     /codon_start=1
                     /product="replication protein"
                     /protein_id="AAA46929.1"
                     /db_xref="GI:496195"
                     /translation="MADDSGTENEGSGCTGWFMVEAIVEHTTGTQISEDEEEEVEDSG
                     YDMVDFIDDRHITQNSVEAQALFNRQEADAHYATVQDLKRKYLGSPYVSPISNVANAV
                     ESEISPRLDAIKLTTQPKKVKRRLFETRELTDSGYGYSEVEAATQVEKHGDPENGGDG
                     QERDTGRDIEGEGVEHREAEAVDDSTREHADTSGILELLKCKDIRSTLHGKFKDCFGL
                     SFVDLIRPFKSDRTTCADWVVAGFGIHHSIADAFQKLIEPLSLYAHIQWLTNAWGMVL
                     LVLIRFKVNKSRCTVARTLGTLLNIPENHMLIEPPKIQSGVRALYWFRTGISNASTVI
                     GEAPEWITRQTVIEHSLADSQFKLTEMVQWAYDNDICEESEIAFEYAQRGDFDSNARA
                     FLNSNMQAKYVKDCAIMCRHYKHAEMKKMSIKQWIKYRGTKVDSVGNWKPIVQFLRHQ
                     NIEFIPFLSKLKLWLHGTPKKNCIAIVGPPDTGKSCFCMSLIKFLGGTVISYVNSCSH
                     FWLQPLTDAKVALLDDATQPCWTYMDTYMRNLLDGNPMSIDRKHRALTLIKCPPLLVT
                     SNIDISKEEKYKYLHSRVTTFTFPNPFPFDRNGNAVYELSDANWKCFFERLSSSLDIE
     gene            DSEDEEDGSNSQAFRCVPGSVVRTL"
                     2723. .3826
     CDS             /gene="E2"
                     2723. .3826
                     /gene="E2"
                     /note="2723 is position of first start codon in ORF E2;
                     putative"
                     /codon_start=1
                     /product="regulatory protein"
                     /protein_id="AAA46930.1"
                     /db_xref="GI:496196"
                     /translation="MEAIAKRLDACQDQLLELYEENSIDIHKHIMHWKCIRLESVLLH
                     KAKQMGLSHIGLQVVPPLTVSETKGHNAIEMQMHLESLAKTQYGVEPWTLQDTSYEMW
                     LTPPKRCFKKQGNTVEVKFDGCEDNVMEYVVWTHIYLQDNDSWVKVTSSVDAKGIYYT
                     CGQFKTYYVNFNKEAQKYGSTNHWEVCYGSTVICSPASVSSTVREVSIAEPTTYTPAQ
                     TTAPTVSACTTEDGVSAPPRKRARGPSTNNTLCVANIRSVDSTINNIVTDNYNKHQRR
                     NNCHSAATPIVQLQGDSNCLKCFRYRLNDKYKHLFELASSTWHWASPEAPHKNAIVTL
     gene            TYSSEEQRQQFLNSVKIPPTIRHKVGFMSLHLL"
                     3255. .3581
     CDS             /gene="E4"
                     3255. .3581
                     /gene="E4"
                     /note="3255 is position of first start codon in ORF E4;
                     putative"
                     /codon_start=1
                     /protein_id="AAA46931.1"
                     /db_xref="GI:496197"
                     /translation="MVVPIIGKYVMAAQLYVLLHLYLALYEKYPLLNLLHTPPHRPPP
                     LQCPPAPRKTACRRRLGSEHVDRPLTTPCVWPTSDPWTVQSTTSSLTITTSTKEGTTV
     gene            TVQLRL"
                     3871. .4146
     CDS             /gene="E5A"
                     3871. .4146
                     /gene="E5A"
                     /note="3871 is position of first start codon in ORFE5A;
                     putative"
                     /codon_start=1
                     /protein_id="AAA46932.1"
                     /db_xref="GI:496198"
                     /translation="MEVVPVQIAAATTTTLILPVVIAFAVCILSIVLIILISDFVVYT
     gene            SVLVLTLLLYLLLWLLLTTPLQFFLLTLCVCYFPAFYIHIYIVQTQQ"
                     4146. .4370
     CDS             /gene="E5B"
                     4146. .4370
                     /gene="E5B"
                     /note="4146 is position of first start codon in ORF E5B;
                     putative"
                     /codon_start=1
                     /protein_id="AAA46933.1"
                     /db_xref="GI:496199"
                     /translation="MVMLTCHLNDGDTWLFLWLFTAFVVAVLGLLLLHYRAVHGTEKT
                     KCAKCKSNRNTTVDYVYMSHGDNGDYVYMN"
     polyA_signal    4371. .4376
     gene            /note="putative"
                     4417. .5784
     CDS             /gene="L2"
                     4417. .5784
                     /gene="L2"
                     /note="4417 is position of first start codon in ORF L2;
                     putative"
                     /codon_start=1
                     /product="minor capsid protein"
                     /protein_id="AAA46934.1"
                     /db_xref="GI:496200"
                     /translation="MKPRARRRKRASATQLYQTCKATGTCPPDVIPKVEHTTIADQIL
                     KWGSLGVFFGGLGIGTGAGSGGRAGYIPLGSSPKPAITGGPAARPPVLVEPVAPSDPS
                     IVSLIEESAIINAGAPEVVPPTQGGFTITSSESTTPAILDVSVTNHTTTSVFQNPLFT
                     EPSVIQPQPPVEASGHILISAPTITSQHVEDIPLDTFVVSSSDSGPTSSTPLPRAFPR
                     PRVGLYSRALQQVQVTDPAFLSTPQRLVTYDNPVYEGEDVSLQFTHESIHNAPDEAFM
                     DIIRLHRPAITSRRGLVRFSRIGQRGSMYTRSGQHIGARIHYFQDISPVTQAAEEIEL
                     HPLVAAENDTFDIYAEPFDPIPDPVQHSVTQSYLTSTPNTLSQSWGNTTVPLSIPSDW
                     FVQSGPDITFPTASMGTPFSPVTPALPTGPVFITGSDFYLHPTWYFARRRRKRIPLFF
                     TDVAA"
     polyA_signal    4545. .4551
                     /gene="L2"
     gene            /note="putative"
                     5771. .7276
     CDS             /gene="L1"
                     5771. .7276
                     /gene="L1"
                     /note="5771 is position of first start codon in ORF L1;
                     putative"
                     /codon_start=1
                     /product="major capsid protein"
                     /protein_id="AAA46935.1"
                     /db_xref="GI:496201"
                     /translation="MWRPSDSTVYVPPPNPVSKVVATDAYVKRTNIFYHASSSRLLAV
                     GHPYYSIKKVNKTVVPKVSGYQYRVFKVVLPDPNKFALPDSSLFDPTTQRLVWACTGL
                     EVGRGQPLGVGVSGHPLLNKYDDVENSGGYGGNPGQDNRVNVGMDYKQTQLCMVGCAP
                     PLGEHWGKGTQCSNTSVQNGDCPPLELITSVIQDGDMVDTGFGAMNFADLQTNKSDVP
                     LDICGTVCKYPDYLQMAADPYGDRLFFYLRKEQMFARHFFNRAGTVGEPVPDDLLVKG
                     GNNRSSVASSIYYVHTPSGSLVSSEAQLFNKPYWLQKAQGHNNGICWGNHLFVTVVDTT
                     RSTNMTLCASVSKSATYTNSDYKEYMRHVEEFDLQFIFQLCSITLSAEVMAYIHTMNP
                     SVLEDWNFGLSPPPNGTLEDTYRYVQSQAITCQKPTPEKEKQDPYKDMSFWEVNLKEK
     5'UTR           FSSELDQFPLGRKFLLQSGYRGRTSARTGIKRPAVSKPSTAPKRKRTKTKK"
                     join(7277. .7931,1. .101)
                     /standard_name="LCR"
     repeat_region   /function="regulatory region"
                     7320. .7360
                     /standard_name="direct repeats"
     repeat_region   /note="putative"
                     7339. .7374
                     /standard_name="direct repeats"
     repeat_region   /note="putative"
                     7374. .7403
     polyA_signal    /note="putative"
                     7457. .7462
     protein_bind    /note="putative"
                     7592. .7603
                     /note="putative"
                     /function="gene transcription"
     polyA_signal    /bound_moiety="E2"
                     7748. .7753
     protein_bind    /note="putative"
                     7890. .7901
                     /note="putative"
                     /function="gene transcription"
                     /bound_moiety="E2"
BASE COUNT      2406 a    1519 c    1736 g    2270 t
ORIGIN      4557 bp upstream of HindIII site.

PPH11  Length: 7931  November 28, 2001 14:10  Type: N  Check: 3689
```

```
Initial Score    =     12  Optimized Score  =     13  Significance =  1.94
Residue Identity =     60% Matches          =     14  Mismatches   =     8
Gaps             =      1  Conservative Substitutions               =     0

                                                      X       10        20
                                                      TCATCTCGAACGTTCACAGT-C
                                                       ||||     |||  ||||| |
    TAGGAAGCGAGCACGTGGACCGTCCACTAACAACACCCTGTGTGTGGCCAACATCAGATCCGTGGACAGTAC
        3440      3450      3460      3470      3480  X   3490      3500

    X
    A
    |
    AATCAACAACATCGTCACTGACAATTACAACAAGCACCAAAGAAGGAACAA
    X 3510      3520      3530      3540      3550


2. SEQ1' (1-22)
   af131950     TOIG of: af131950  check: 557  from: 1  to: 7812

  TOIG of: af131950  check: 557  from: 1  to: 7812

LOCUS       AF131950     7812 bp    DNA             VRL       07-FEB-2001
DEFINITION  Human papillomavirus candHPV85, complete genome.
ACCESSION   AF131950
VERSION     AF131950.1  GI:4574720
KEYWORDS    .
SOURCE      Human papillomavirus candHPV85
  ORGANISM  Human papillomavirus candHPV85
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 7812)
  AUTHORS   Chow,V.T.K. and Leong,P.W.F.
  TITLE     Complete nucleotide sequence, genomic organization and phylogenetic
            analysis of a novel genital human papillomavirus type, HLT7474-S
  JOURNAL   J. Gen. Virol. 80 (Pt 11), 2923-2929 (1999)
  MEDLINE   20047972
   PUBMED   10580054
REFERENCE   2  (bases 1 to 7812)
  AUTHORS   Chow,V.T.K. and Leong,W.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-FEB-1999) Department of Microbiology, National
            University of Singapore, 10 Kent Ridge Crescent, Singapore 119260,
            Singapore
FEATURES             Location/Qualifiers
     source          1. .7812
                     /organism="Human papillomavirus candHPV85"
                     /db_xref="taxon:151757"
                     /note="isolated from scraped uterine cervical cells from a
                     female sex worker; overlapping PCR products"
     protein_bind    40. .51
                     /note="putative"
                     /bound_moiety="E2 protein"
                     /function="transcriptional regulation"
     protein_bind    56. .67
                     /note="putative"
                     /bound_moiety="E2 protein"
                     /function="transcriptional regulation"
     TATA_signal     71. .79
                     /note="putative"
     gene            105. .578
                     /gene="E6"
     CDS             105. .578
                     /gene="E6"
                     /codon_start=1
                     /product="putative transforming protein E6"
                     /protein_id="AAD24181.1"
                     /db_xref="GI:4574721"
                     /translation="MAEFGNPATRPYKLPDLCNTLDTSLQDIEISCVYCKSVLQRTEV
                     YEFAFADLFVVYRDGIPYAACQNCLMFYSKIRELRYYSDSVYGETLEKLTNSNIYDIL
                     IRCLRCQKPLCPAEKLKHLNEKRRFHKIAGKYRGQCRRCMTRAQEQQGSRRETQV"
     gene            587. .913
                     /gene="E7"
     CDS             587. .913
                     /gene="E7"
                     /codon_start=1
                     /product="putative transforming protein E7"
                     /protein_id="AAD24182.1"
                     /db_xref="GI:4574722"
                     /translation="MHGPKPTVHDIVLDLEPYNEVQEVDLYCYEELNNSEEEIDEPDN
                     AINHRQPLLARREELQRHTICCVCCKCEASLQLVVESSAADLRDLQQLFLGTLSFLCP
                     LCAVLR"
     gene            920. .2866
                     /gene="E1"
     CDS             920. .2866
                     /gene="E1"
                     /codon_start=1
                     /product="putative replication protein E1"
                     /protein_id="AAD24183.1"
                     /db_xref="GI:4574723"
                     /translation="MADPEGTDGDGTGCNGWFFVQAIVDKKTGDTISEDEDEDGTDTG
                     SDLVDFIDTSNTNYMQADREAAQALLHAQEVETDTKLLHALKRKYGAHSTENSPCRDT
                     ASIHSNLSSPLQEISLNSYNNTAKRRLCSVPDSGYGNTQVETLQTQVTLDTNVFGDGK
                     NGDGLNSEACSTDNEMDIENQNPNSPMTQIVSLLKVNNKKAAILAKFKETYGLSFTDL
                     VRTFKSDKTTCTDWVAAICGVNPNIAEGFKTLIQPYVLYAHIQCMDCSWGVFILALLR
                     YKCGKNRLTVAKGLSTLLHVPDTHMLIEPPKLRSSCAALYWYRTGISNISEVTGDTPE
                     WIQRQTIIQHGIDDSVFDLSEMIQWAFDNDYIDESDIAYEYAQLADCNSNAAAFLKSN
                     CQAKYLRDCAVMCRHYKRAQRKQMNMSQWISYRCDKIDDGGDWKPIVQFLRFQGIEFI
                     TFLRAFKDFLKGTPKKNCIVIYGPANTGKSYFCMSLIQFLHGTVLSFVNSNSHFWLEP
                     LTDTKIAMVDDATPTCWSYFDNYMRNALDGNPISIDRKHKHLIQMKCPPMLITSNTNP
                     ATDDRWPYLRSRVTVFTFPHTFPFDSNGNPVYDINDKNWKCFFKRTWSRLDLHQEEED
                     TENDGCPIPTFKCVTGKDSRTLRD"
     gene            2799. .3932
                     /gene="E2"
     CDS             2799. .3932
                     /gene="E2"
                     /codon_start=1
                     /product="putative regulatory protein E2"
                     /protein_id="AAD24184.1"
                     /db_xref="GI:4574724"
                     /translation="MMDALSQRLNVLQEKILEHYETDSTDICDQIDYWKCVRLENAIY
                     YAAREHGLKTINHHVVPTFQISKSKAHEAIELQMALESLAKSEFKNELWTLQDTCQEL
                     YQTPPQQCFKKQGQTVEVRYDGDKDNTMHYTSWDYIYYVTEGDKWCKTKGYVNYCGLY
                     YIKEGQQTYYVQFKCDAQQYGQSGKWEVWYNGKKIECVNTCDKIIQCSESMYSTCDET
                     VSATAIARELQHPTTPYTEATTVCTQKSGGSAPTRNPFRHCGFTETSEVDGLSVDHLN
                     NPLLSAPAGNNGFRKHNSGDTTPIVHLKGDKNRLKCLRYRLQKFHNLYTQISCTWHWI
                     GGKETTRTGILTLAYANETQRQKFLDVVKIPNTVNVSLGYMTL"
     gene            <3412. .3699
                     /gene="E4"
     CDS             <3412. .3699
                     /gene="E4"
                     /note="start codon absent"
                     /codon_start=1
                     /product="putative protein E4"
                     /protein_id="AAD24185.1"
                     /db_xref="GI:4574725"
                     /translation="YNVLNLCTVPVTRQYPLLQLLENYNTPPHRIPKPPPCAPKRAGV
                     RRRLETHSDIVDSQKQVKSTDCPWTTSTTPCLVHLQATTGSGSTIVVTLRL"
     gene            3940. .4161
                     /gene="E5"
     CDS             3940. .4161
                     /gene="E5"
                     /codon_start=1
                     /product="putative protein E5"
                     /protein_id="AAD24186.1"
                     /db_xref="GI:4574726"
                     /translation="MLIAVFLLCFCVCLYVCCTIPLFPATNVCAHAWLLVFLLIVFIT
                     TPFTAFIVYIFFFILPMFLLHLHSVHTFD"
     gene            4226. .5629
                     /gene="L2"
     CDS             4226. .5629
                     /gene="L2"
                     /codon_start=1
                     /product="putative minor capsid protein L2"
                     /protein_id="AAD24187.1"
                     /db_xref="GI:4574727"
```

```
                     /translation="MVSHRAARRKRASATDLYKTCKQSGTCPPDVINKVEGTTLADKL
                     LQWTSLGIFLGGLGIGTGSGTGGRTGYIPLGGRPNTVVDVSPARPPVVIESVGPSDPS
                     IVTLVEESSIVTSGAPVPTFTGTSGFEITSSATTTPAVLDITPASGSVQLSSTSFTNP
                     AFTDPSVIEVPQTGEVSGDIFITTPTSGTHGYEEIPMHTFATQGRGTEPISSTPIPGV
                     RRVAGPRLYSQAYQQVKITNSDFISRPSTLVTFTNPAYEPIDTTLTFSPQDVVPDPDF
                     MDIVRLHRPALTSRRGTVRFSRLGKKLTMSTRSGKQIGAQVHYYHDISPISHIGESIE
                     MQPLLPDAAVTADTNGLFDIYADTDIDNNAMLYDRNISDVTQPTTSTISSVSSRYSNT
                     TIPLATSWDVPVHTGPDMTLPTTIPQWPNIVPLLPNNTHSVVLQGTNYYLWPNYYFIF
     gene            KKRKRVPYFLTDGFVAF"
                     5610. .7124
     CDS             /gene="L1"
                     5610. .7124
                     /gene="L1"
                     /codon_start=1
                     /product="putative major capsid protein L1"
                     /protein_id="AAD24188.1"
                     /db_xref="GI:4574728"
                     /translation="MALWRSSDSKVYLPPPSVAKVVNTDDYVTRTSTFYHAGSSRLLT
                     VGHPYYKVTSNGGRKQDIPKVSAYQYRVFRVTLPDPNKFGLPDTNIYNPETQRLVWAC
                     VGMEVGRGQPLGVSLSGHPFYNKLDDTENSHVATSVVTHDTRDNVSVDYKQTQLCIIG
                     CVPAIGEHWAKGTACKPGAVQTGDCPPLELVNTPIEDGDMIDTGYGAMDFSTLQDNKS
                     EVPLDICQSICKYPDYLQMSADAYGDSMFFCLRREQLFARHFWNRGGTIGDAVPETLY
                     IKGTNDRATPGSCIYSPSPSGSMVSSDAQMFNKPYWLHKAQGQNNGICWHNQLFITVV
                     DTTRSTNLTLSTATTNPVPSIYEPSKFKEYTRHVEEYDLQFIFQLCKITLTTDVMSYI
                     HNMDPTILDSWNFGVSPPPSASLVDTYRFLQSSAITCQKDVVVPQKKDPYEKLKFWNV
                     DLKEHFSSDLDQFPLGRKFLLQAGLRPKPTIGPRKRVASTSTATRPSKRKRTAK"
     protein_bind    7456. .7467
                     /note="putative"
                     /bound_moiety="E2 protein"
     misc_feature    /function="transcriptional regulation"
                     7647. .7666
     protein_bind    /note="consensus motif of oncogenic anogenital HPV types"
                     7776. .7787
                     /note="putative"
                     /bound_moiety="E2 protein"
                     /function="transcriptional regulation"
BASE COUNT      2502 a    1391 c    1558 g    2361 t
ORIGIN

AF131950  Length: 7812  November 28, 2001 14:10  Type: N  Check: 557  ..

Initial Score    =     11  Optimized Score  =     13  Significance  =  0.97
Residue Identity =    62%  Matches          =     15  Mismatches    =     7
Gaps             =      2  Conservative Substitutions               =     0

                                                    X        10         20
                                                    TCATCTCGA-ACGT-TCACAGT
                                                    | |  || | |||| |||||
    TGTGCCTGATCCTGATTTTATGGATATTGTTCGTTTACATAGGCCTGCCTTAACATCCAGACGTGGTACAGT
      5040      5050      5060      5070      5080    X 5090      5100

     X
     CA
     |
    TAGATTTAGTAGATTGGGTAAAAAACTAACTATGTCTACTCGCAGTGGTAAA
    5110      5120      5130      5140      5150

3. SEQ1' (1-22)
   pph16        TOIG of: pph16  check: 6074  from: 1  to: 7904

 TOIG of: pph16  check: 6074  from: 1  to: 7904

LOCUS       PPH16          7904 bp    DNA    circular  VRL       18-MAR-1994
DEFINITION  Human papillomavirus type 16 (HPV16), complete genome.
ACCESSION   K02718
VERSION     K02718.1  GI:333031
KEYWORDS    circular; complete genome.
SOURCE      Papilloma virus type 16 DNA, isolated from a human invasive
            cervical carcinoma.
  ORGANISM  Human papillomavirus type 16
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
```

```
REFERENCE   1  (bases 1 to 7904)
  AUTHORS   Seedorf,K., Kraemmer,G., Duerst,M., Suhai,S. and Roewekamp,W.G.
  TITLE     Human papillomavirus type 16 DNA sequence
  JOURNAL   Virology 145, 181-185 (1985)
  MEDLINE   85246220
REFERENCE   2  (sites)
  AUTHORS   Kennedy,I.M., Haddow,J.K. and Clements,J.B.
  TITLE     A negative element in the human poapillomavirus type 16 genome acts
            at the level of late mRNA stability
  JOURNAL   J. Virol. 65, 2093-2097 (1991)
  MEDLINE   91162763
COMMENT     The sense strand of this double-stranded circular genome is shown,
            with a numbering system matching the first 60 bp of HPVa1, HPV6b
            and BPV1. The annotation of sites and features is solely based upon
            homology comparison with these other papillomaviruses. In addition
            to the coding sequences reported below, the authors note open
            reading frames which do not start with 'ATG', but which are found
            in other papillomaviruses. In particular, a second portion of the
            E1 gene may be located out to base 2813 (the E1 protein is thought
            to be generally involved in DNA replication).
            A potential 'CAT'-box region is found beginning at base 7895 below,
            and 'TATA' boxes for early and late transcripts may be located at
            17, 65 and 4289. Potential polyadenylation signals are at bases
            4213 and 7260.
            HPV16, in comparison to HPV types 6 and 11, is more often
            associated with malignant genital cancers in humans.
FEATURES            Location/Qualifiers
     source         1. .7904
                    /organism="Human papillomavirus type 16"
                    /db_xref="taxon:10581"
     TATA_signal    17. .23
     TATA_signal    65. .71
     gene           83. .559
                    /gene="E6"
     CDS            83. .559
                    /gene="E6"
                    /note="E6 ORF from 65 to 559; putative"
                    /codon_start=1
                    /product="transforming protein"
                    /protein_id="AAA46939.1"
                    /db_xref="GI:333032"
                    /translation="MHQKRTAMFQDPQERPRKLPQLCTELQTTIHDIILECVYCKQQL
                    LRREVYDFAFRDLCIVYRDGNPYAVCDKCLKFYSKISEYRHYCYSLYGTTLEQQYNKP
     gene           LCDLLIRCINCQKPLCPEEKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL"
                    562. .858
     CDS            /gene="E7"
                    562. .858
                    /gene="E7"
                    /note="E7 ORF from 544 to 858; putative"
                    /codon_start=1
                    /product="transforming protein"
                    /protein_id="AAA46940.1"
                    /db_xref="GI:333033"
                    /translation="MHGDTPTLHEYMLDLQPETTDLYCYEQLNDSSEEEDEIDGPAGQ
     gene           AEPDRAHYNIVTFCCKCDSTLRLCVQSTHVDIRTLEDLLMGTLGIVCPICSQKP"
                    join(865. .1140,1140. .2813)
                    /gene="E1"
     CDS            join(865. .1140,1140. .2813)
                    /gene="E1"
                    /note="E1 interrupted ORF from 859 to 2813; putative"
                    /codon_start=1
                    /product="replication protein"
                    /protein_id="AAA46936.1"
                    /db_xref="GI:459912"
                    /translation="MADPAGTNGEEGTGCNGWFYVEAVVEKKTGDAISDDENENDSDT
                    GEDLVDFIVNDNDYLTQAETETAHALFTAQEAKQHRDAVQVLKRKYLVSPLSDISGCV
                    DNNISPRLKAICIEKQSRAAKRRLFESEDSGYGNTEVETQQMLQVEGRHETETPCSQY
                    SGGSGGGCSQYSSGSGGEGVSERHTICQTPLTNILNVLKTSNAKAAMLAKFKELYGVS
                    FSELVRPFKSNKSTCCDWCIAAFGLTPSIADSIKTLLQQYCLYLHIQSLACSWGMVVL
                    LLVRYKCGKNRETIEKLLSKLLCVSPMCMMIEPPKLRSTAAALYWYKTGISNISEVYG
                    DTPEWIQRQTVLQHSFNDCTFELSQMVQWAYDNDIVDDSEIAYKYAQLADTNSNASAF
                    LKSNSQAKIVKDCATMCRHYKRAEKKQMSMSQWIKYRCDRVDDGGDWKQIVMFLRYQG
                    VEFMSFLTALKRFLQGIPKKNCILLYGAANTGKSLFGMSLMKFLQGSVICFVNSKSHF
```

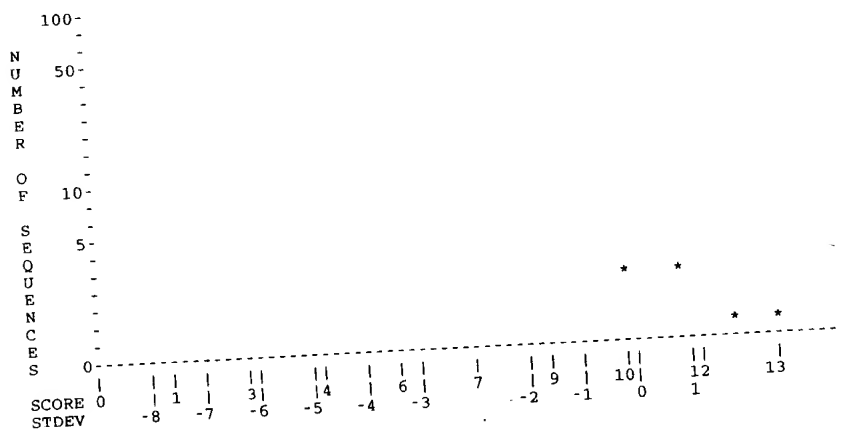
```
                WLQPLADAKIGMLDDATVPCWNYIDDNLRNALDGNLVSMDVKHRPLVQLKCPPLLITS
                NINAGTDSRWPYLHNRLVVFTFPNEFPFDENGNPVYELNDKNWKSFFSRTWSRLSLHE
                DEDKENDGDSLPTFKCVSGQNTNTL"
gene            2755. .3852
                /gene="E2"
CDS             2755. .3852
                /gene="E2"
                /note="E2 ORF from 2725 to 3852; putative"
                /codon_start=1
                /product="regulatory protein"
                /protein_id="AAA46941.1"
                /db_xref="GI:333035"
                /translation="METLCQRLNVCQDKILTHYENDSTDLRDHIDYWKHMRLECAIYY
                KAREMGFKHINHQVVPTLAVSKNKALQAIELQLTLETIYNSQYSNEKWTLQDVSLEVY
                LTAPTGCIKKHGYTVEVQFDGDICNTMHYTNWTHIYICEEASVTVVEGQVDYYGLYYV
                HEGIRTYFVQFKDDAEKYSKNKVWEVHAGGQVILCPTSVFSSNEVSSPEIIRQHLANH
                PAATHTKAVALGTEETQTTIQRPRSEPDTGNPCHTTKLLHRDSVDSAPILTAFNSSHK
                GRINCNSNTTPIVHLKGDANTLKCLRYRFKKHCTLYTAVSSTWHWTGHNVKHKSAIVT
                LTYDSEWQRDQFLSQVKIPKTITVSTGFMSI"
gene            3332. .3619
                /gene="E4"
CDS             <3332. .3619
                /gene="E4"
                /note="E4 ORF from 3332 to 3619; putative"
                /codon_start=1
                /protein_id="AAA46937.1"
                /db_xref="GI:459913"
                /translation="YYVLHLCLAATKYPLLKLLGSTWPTTPPRPIPKPSPWAPKKHRR
                LSSDQDQSQTPETPATPLSCCTETQWTVLQSSLHLTAHTKDGLTVIVTLHP"
gene            3863. .4099
                /gene="E5"
CDS             <3863. .4099
                /gene="E5"
                /note="E5 ORF from 3863 to 4099; putative"
                /codon_start=1
                /protein_id="AAA46938.1"
                /db_xref="GI:459914"
                /translation="YCIHNITGVLFALLCVLLCVCLLIRPLLLSVSTYTSLIILVLLL
                WITAASAFRCFIVYIIFVYIPLFLIHTHARFLIT"
polyA_signal    4213. .4218
                /note="putative"
gene            4235. .5656
                /gene="L2"
CDS             4235. .5656
                /gene="L2"
                /note="L2 ORF from 4133 to 5656; putative"
                /codon_start=1
                /product="minor capsid protein"
                /protein_id="AAA46942.1"
                /db_xref="GI:333036"
                /translation="MRHKRSAKRTKRASATQLYKTCKQAGTCPPDIIPKVEGKTIAEQ
                ILQYGSMGVFFGGLGIGTGSGTGGRTGYIPLGTRPPTATDTLAPVRPPLTVDPVGPSD
                PSIVSLVEETSFIDAGAPTSVPSIPPDVSGFSITTSTDTTPAILDINNTVTTVTTHNN
                PTFTDPSVLQPPTPAETGGHFTLSSSTISTHNYEEIPMDTFIVSTNPNTVTSSTPIPG
                SRPVARLGLYSRTTQQVKVVDPAFVTTPTKLITYDNPAYEGIDVDNTLYFSSNDNSIN
                IAPDPDFLDIVALHRPALTSRRTGIRYSRIGNKQTLRTRSGKSIGAKVHYYYDLSTID
                PAEEIELQTITPSTYTTTSHAASPTSINNGLYDIYADDFITDTSTTPVPSVPSTSLSG
                YIPANTTIPFGGAYNIPLVSGPDIPINITDQAPSLIPIVPGSPQYTIIADAGDFYLHP
                SYYMLRKRRKRLPYFFSDVSLAA"
TATA_signal     4289. .4295
                /gene="L2"
gene            5559. .7154
                /gene="L1"
CDS             5559. .7154
                /gene="L1"
                /note="L1 ORF from 5526 to 7154; putative"
                /codon_start=1
                /product="major capsid protein"
                /protein_id="AAA46943.1"
                /db_xref="GI:333037"
                /translation="MQVTFIYILVITCYENDVNVYHIFFQMSLWLPSEATVYLPPVPV
                SKVVSTDEYVARTNIYYHAGTSRLLAVGHPYFPIKKPNNNKILVPKVSGLQYRVFRIH
                LPDPNKFGFPDTSFYNPDTQRLVWACVGVEVGRGQPLGVGISGHPLLNKLDDTENASA
                YAANAGVDNRECISMDYKQTQLCLIGCKPPIGEHWGKGSPCTNVAVNPGDCPPLELIN
                TVIQDGDMVHTGFGAMDFTTLQANKSEVPLDICTSICKYPDYIKMVSEPYGDSLFFYL
                RREQMFVRHLFNRAGTVGENVPDDLYIKGSGSTANLASSNYFPTPSGSMVTSDAQIFN
                KPYWLQRAQGHNNGICWGNQLFVTVVDTTRSTNMSLCAAISTSETTYKNTNFKEYLRH
                GEEYDLQFIFQLCKITLTADVMTYIHSMNSTILEDWNFGLQPPPGGTLEDTYRFVTQA
                IACQKHTPPAPKEDDPLKKYTFWEVNLKEKFSADLDQFPLGRKFLLQAGLKAKPKFTL
                GKRKATPTTSSTSTTAKRKKRKL"
polyA_signal    7260. .7265
BASE COUNT      2601 a   1377 c   1509 g   2417 t
ORIGIN      Unreported.

PPH16  Length: 7904  November 28, 2001 14:10  Type: N  Check: 6074  ..

Initial Score  =   11  Optimized Score  =   11  Significance =  0.97
Residue Identity =  53%  Matches          =   14  Mismatches   =     8
Gaps           =    4  Conservative Substitutions             =     0

                                                        X         10
                                                        TCATCTCGA--ACGTTCA--CA
                                                         ||||  ||  || | ||

    TAAGCACGGATGAATATGTTGCACGCACAAACATATATTATCATGCAGGAACATCCAGACTACTTGCAGTTG
        5710      5720      5730      5740      5750      5760      5770

    20 X
    GTCA
    | ||
    GACATCCCTATTTTCCTATTAAAAAACCTAACAATAACAAAATATTAGTTCCTA
        5780      5790      5800      5810      5820


4. SEQ1' (1-22)
   pph31a       TOIG of: pph31a  check: 5866  from: 1  to: 7912

   TOIG of: pph31a  check: 5866  from: 1  to: 7912

LOCUS       PPH31A       7912 bp    DNA   circular  VRL       18-MAR-1994
DEFINITION  Human papillomavirus type 31 (HPV-31) complete genome.
ACCESSION   J04353
VERSION     J04353.1  GI:333048
KEYWORDS    complete genome.
SOURCE      Human papillomavirus type 31 DNA.
  ORGANISM  Human papillomavirus type 31
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 7912)
  AUTHORS   Goldsborough,M.D., DiSelvestre,D., Temple,G.F. and Lorincz,A.T.
  TITLE     Nucleotide sequence of human papillomavirus type 31: A cervical
            neoplasia associated virus
  JOURNAL   Virology 171, 306-311 (1989)
  MEDLINE   89299478
COMMENT     Draft entry and computer-readable copy of sequence [1] kindly
            submitted by M.D.Goldsborough, 05-JUL-1989.

FEATURES             Location/Qualifiers
     source          1. .7912
                     /organism="Human papillomavirus type 31"
                     /db_xref="taxon:10585"
     TATA_signal     19. .24
     TATA_signal     69. .74
     gene            108. .557
                     /gene="E6"
     CDS             108. .557
                     /gene="E6"
                     /note="ORF E6 from bp 39 to 557"
                     /codon_start=1
                     /product="transforming protein"
                     /protein_id="AAA46950.1"
                     /db_xref="GI:459916"
                     /translation="MFKNPAERPRKLHELSSALEIPYDELRLNCVYCKGQLTETEVLD
                     FAFTDLTIVYRDDTPHGVCTKCLRFYSKVSEFRWYRYSVYGTTLEKLTNKGICDLLIR
                     CITCQRPLCPEEKQRHLDKKKRFHNIGGRWTGRCIACWRRPRTETQV"
     misc_feature    228. .236
                     /gene="E6"
                     /standard_name="Splice donor"
```

```
misc_feature    403. .414
                /gene="E6"
                /standard_name="Splice acceptor"
gene            560. .856
CDS             /gene="E7"
                560. .856
                /gene="E7"
                /note="ORF E7 from bp 545 to 856"
                /codon_start=1
                /product="transforming protein"
                /protein_id="AAA46951.1"
                /db_xref="GI:459917"
                /translation="MRGETPTLQDYVLDLQPEATDLHCYEQLPDSSDEEDVIDSPAGQ
gene            AEPDTSNYNIVTFCCQCKSTLRLCVQSTQVDIRILQELLMGSFGIVCPNCSTRL"
CDS             862. .2751
                /gene="E1"
                862. .2751
                /gene="E1"
                /note="ORF E1 from bp 850 to bp 2751"
                /codon_start=1
                /product="replication protein"
                /protein_id="AAA46952.1"
                /db_xref="GI:459918"
                /translation="MADPAGTDGEGTGCNGWFYVEAVIDRQTGDNISEDENEDSSDTG
                EDMVDFIDNCNVYNNQAEAETAQALFHAQEAEEHAEAVQVLKRKYVGSPLSDISSCVD
                YNISPRLKAICIENNSKTAKRRLFELPDSGYGNTEVETQQMVQVEEQQTTLSCNGSDG
                THSERENETPTRNILQVLKTSNGKAAMLGKFKELYGVSFMELIRPFQSNKSTCTDWCV
                AAFGVTGTVAEGFKTLLQPYCLYCHLQSLACSWGMVMLMLVRFKCAKNRITIEKLLEK
                LLCISTNCMLIQPPKLRSTAAALYWYRTGMSNISDVYGETPEWIERQTVLQHSFNDTT
                FDLSQMVQWAYDNDVMDDSEIAYKYAQLADSDSNACAFLKSNSQAKIVKDCGTMCRHY
                KRAEKRQMSMGQWIKSRCDKVSDEGDWRDIVKFLRYQQIEFVSFLSALKLFLKGVPKK
                NCILIHGAPNTGKSYFGMSLISFLQGCIISYANSKSHFWLQPLADAKIGMLDDATTPC
                WHYIDNYLRNALDGNPVSIDVKHKALMQLKCPPLLITSNINAGKDDRWPYLHSRLVVF
                TFPNPFPFDKNGNPVYELSDKNWKSFFSRTWCRLNLHEEEDKENDGDSFSTFKCVSGQ
                NIRTL"
gene            2693. .3811
CDS             /gene="E2"
                2693. .3811
                /gene="E2"
                /note="ORF E2 from bp 2663 to 3811"
                /codon_start=1
                /product="regulatory protein"
                /protein_id="AAA46953.1"
                /db_xref="GI:459919"
                /translation="METLSQRLNVCQDKILEHYENDSKRLCDHIDYWKHIRLECVLMY
                KAREMGIHSINHQVVPALSVSKAKALQAIELQMMLETLNNTEYKNEDWTMQQTSLELY
                LTAPTGCLKKHGYTVEVQFDGDVHNTMHYTNWKFIYLCIDGQCTVVEGQVNCKGIYYV
                HEGHITYFVNFTEEAKKYGTGKKWEVHAGGQVIVFPESVFSSDEISFAGIVTKLPTAN
                NTTTSNSKTCALGTSEGVRRATTSTKRPRTEPEHRNTHHPNKLLRGDSVDSVNCGVIS
                AAACTNQTRAVSCPATTPIIHLKGDANILKCLRYRLSKYKQLYEQVSSTWHWTCTDGK
gene            HKNAIVTLTYISTSQRDDFLNTVKIPNTVSVSTGYMTI"
CDS             3270. .3578
                /gene="E4"
                <3270. .3578
                /gene="E4"
                /note="ORF E4 from bp 3270 to 3578"
                /codon_start=1
                /protein_id="AAA46949.1"
                /db_xref="GI:459915"
                /translation="LFFLNLYLAVTKYPLLGLLQSYQQPTTPPHRIPKPAPWAPVKVC
                GGRRRLLSDQEQSQSTETPTTPTSCCEATPWTVSTVGLSVQLHAQTKQGLSVVLQLHL
gene            "
CDS             3816. .4070
                /gene="E5"
                3816. .4070
                /gene="E5"
                /note="ORF E5 from bp 3804 to 4070"
                /codon_start=1
                /protein_id="AAA46954.1"
                /db_xref="GI:459920"
                /translation="MIELNISTVSIVLCFLLCFCVLLFVCLVIRPLVLSVSVYATLLL
repeat_region   LIVILWVIATSPLRCFCIYVVFIYIPLFVIHTHASFLSQQ"
                4099. .4134

polyA_signal    4138. .4143
                /note="putative"
repeat_region   4143. .4158
gene            4171. .5571
                /gene="L2"
CDS             4171. .5571
                /gene="L2"
                /note="ORF L2 from bp 4060 to bp 5571"
                /codon_start=1
                /product="minor capsid protein"
                /protein_id="AAA46955.1"
                /db_xref="GI:459921"
                /translation="MRSKRSTKRTKRASATQLYQTCKAAGTCPSDVIPKIEHTTIADQ
                ILRYGSMGVFFGGLGIGSGSGTGGRTGYVPLSTRPSTVSEASIPIRPPVSIDPVGPLD
                PSIVSLVEESGIVDVGAPAPIPHPPTTSGFDIATTADTTPAILDVTSVSTHENPTFTD
                PSVLQPPTPAETSGHLLLSSSSISTHNYEEIPMDTFIVSTNNENITSSTPIPGVRRPA
                RLGLYSKATQQVKVIDPTFLSAPKQLITYENPAYETVNAEESLYFSNTSHNIAPDPDF
                LDIIALHRPALTSRRNTVRYSRLGNKQTLRTRSGATIGARVHYYYDISSINPAGESIE
                MQPLGASATTTSTLNDGLYDIYADTDFTVDTPATHNVSPSTAVQSTSAVSAYVPTNTT
                VPLSTGFDIPIFSGPDVPIEHAPTQVFPFPLAPTTPQVSIFVDGGDFYLHPSYYMLKR
                RRKRVSYFFTDVSVAA"
gene            5552. .7066
CDS             /gene="L1"
                5552. .7066
                /gene="L1"
                /note="ORF L1 from bp 5516 to  7066"
                /codon_start=1
                /product="major capsid protein"
                /protein_id="AAA46956.1"
                /db_xref="GI:459922"
                /translation="MSLWRPSEATVYLPPVPVSKVVSTDEYVTRTNIYYHAGSARLLT
                VGHPYYSIPKSDNPKKIVVPKVSGLQYRVFRVRLPDPNKFGFPDTSFYNPETQRLVWA
                CVGLEVGRGQPLGVGISGHPLLNKFDDTENSNRYAGGPGTDNRECISMDYKQTQLCLL
                GCKPPIGEHWGKGSPCSNNAITPGDCPPLELKNSVIQDGDMVDTGFGAMDFTALQDTK
                SNVPLDICNSICKYPDYLKMVAEPYGDTLFFYLRREQMFVRHFFNRSGTVGESVPTDL
                YIKGSGSTATLANSTYFPTPSGSMVTSDAQIFNKPYWMQRAQGHNNGICWGNQLFVTV
                VDTTRSTNMSVCAAIANSDTTFKSSNFKEYLRHGEEFDLQFIFQLCKITLSADIMTYI
                HSMNPAILEDWNFGLTTPPSGSLEDTYRFVTSQAITCQKTAPQKPKEDPFKDYVFWEV
                NLKEKFSADLDQFPLGRKFLLQAGYRARPKFKAGKRSAPSASTTTPAKRKKTKK"
polyA_signal    7227. .7231
                /note="putative"
repeat_region   7291. .7302
repeat_region   7314. .7333
                /note="putative"
protein_bind    7406. .7420
                /standard_name="glucocorticoid responsive element"
                /bound_moiety="hormone receptor"
protein_bind    7477. .7488
                /function="gene transcription"
                /bound_moiety="E2"
enhancer        7542. .7549
                /standard_name="keratinocyte-dependant enhancer"
protein_bind    7868. .7879
                /function="gene transcription"
                /bound_moiety="E2"
BASE COUNT      2528 a   1364 c   1572 g   2448 t
ORIGIN

PPH31A  Length: 7912  November 28, 2001 14:10  Type: N  Check: 5866  ..

Initial Score    =    11  Optimized Score  =    14  Significance  =  0.97
Residue Identity =   63%  Matches          =    14  Mismatches    =     8
Gaps             =     0  Conservative Substitutions              =     0

                                                    X       10        20 X
                                                    TCATCTCGAACGTTCACAGTCA
                                                    |||| |   || ||||   |||
TTTACTGATCCATCTGTATTGCAGCCTCCTACACCTGCAGAAACATCAGGTCATTTACTACTTTCATCATCA
    4650      4660      4670      4680      4690 X    4700      4710  X

TCTATTAGCACACATAATTATGAGGAAATACCTATGGATACATTTATTGT
   4720      4730      4740      4750      4760
```

```
5. SEQ1' (1-22)
   a12360      TOIG of: a12360  check: 1580  from: 1  to: 7909

TOIG of: a12360  check: 1580  from: 1  to: 7909

LOCUS       A12360       7909 bp    DNA              PAT       12-DEC-1993
DEFINITION  complete nucleotide sequence of HPV-33.
ACCESSION   A12360
VERSION     A12360.1  GI:492936
KEYWORDS    .
SOURCE      Human papillomavirus type 33.
  ORGANISM  Human papillomavirus type 33
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 7909)
  AUTHORS   .
  JOURNAL   Patent: WO 8705630-A 1 24-SEP-1987;
FEATURES             Location/Qualifiers
     source          1. .7909
                     /organism="Human papillomavirus type 33"
                     /db_xref="taxon:10586"

BASE COUNT     2544 a   1354 c   1535 g   2474 t      2 others
ORIGIN

A12360  Length: 7909  November 28, 2001 14:10  Type: N  Check: 1580  ..

Initial Score     =     10  Optimized Score  =     11  Significance =  0.00
Residue Identity  =    54%  Matches          =     12  Mismatches   =     9
Gaps              =      1  Conservative Substitutions               =     0

                                                      X       10        20 X
                                                      TCATCTCGAACGTTCACAGTCA
                                                        ||     |||||  |  ||||
   AAGCAAAAATAGTAAAGGACTGTGGAATAATGTGTAGACATTATAAAAAAGCA-GAAAAACGTAAAATGTCA
2050      2060      2070      2080      2090      2100      2110      2120

   ATAGGACAATGGATACAAAGTAGATGTGAAAAAACAAATGATGGAGGAAA
       2130      2140      2150      2160      2170


6. SEQ1' (1-22)
   hpu06714      TOIG of: hpu06714  check: 4862  from: 1  to: 7801

TOIG of: hpu06714  check: 4862  from: 1  to: 7801

LOCUS       HPU06714      7801 bp    DNA              VRL       04-FEB-1997
DEFINITION  Human papillomavirus HPV-1A (3-3), complete genome.
ACCESSION   U06714
VERSION     U06714.1  GI:458704
KEYWORDS    .
SOURCE      Human papillomavirus.
  ORGANISM  Human papillomavirus
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 7801)
  AUTHORS   Danos,O., Katinka,M. and Yaniv,M.
  TITLE     Human papillomavirus 1A complete DNA sequence: a novel type of
            genome organization among Papovaviridae
  JOURNAL   EMBO J. 1, 231-236 (1982)
  MEDLINE   84182467
REFERENCE   2  (bases 1 to 7801)
  AUTHORS   Meissner,J.
  TITLE     Complete nucleotide sequencing of an HPV-1a variant and
            determination of extant errors in the prototype HPV-1a sequence
  JOURNAL   Virus Genes 9 (2), 189-191 (1995)
  MEDLINE   95250312
REFERENCE   3  (bases 1 to 7801)
  AUTHORS   Meissner,J.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-FEB-1994) John D. Meissner, Duke University Medical
            Center, Microbiology, 277 Carl Building, Research Drive, Durham, NC
            27710 USA

FEATURES             Location/Qualifiers
     source          1. .7801
                     /organism="Human papillomavirus"
                     /strain="HPV-1A (3-3)"
                     /db_xref="taxon:10566"
     variation       142
                     /replace="a"
     conflict        1283
                     /citation=[1]
                     /replace="a"
     conflict        2301
                     /citation=[1]
                     /replace="t"
     variation       2825
                     /replace="t"
     conflict        3884. .3886
                     /citation=[1]
                     /replace="aa"
     conflict        4332
                     /citation=[1]
                     /replace="a"
     conflict        4376. .4382
                     /citation=[1]
                     /replace="gagggaa"
     variation       5794
                     /replace="a"
     variation       6905
                     /replace="g"
     variation       7186. .7187
                     /note="15 bp deletion"
                     /citation=[1]
                     /replace="ctagattgtattgctat"
     variation       7560
                     /replace="c"
     variation       7618
                     /replace="g"
     conflict        7677. .7678
                     /citation=[1]
                     /replace="cc"
     variation       7787
                     /replace="g"

BASE COUNT     2389 a   1482 c   1664 g   2266 t
ORIGIN

HPU06714  Length: 7801  November 28, 2001 14:10  Type: N  Check: 4862  ..

Initial Score     =      9  Optimized Score  =     12  Significance = -0.97
Residue Identity  =    65%  Matches          =     15  Mismatches   =     4
Gaps              =      4  Conservative Substitutions               =     0

                                                      X       10        20
                                                      TCATC-TCGAACGTTCACAGTC
                                                      |  || || |||. || ||||
   GGAAGAATTGGTTATACTCCCCTCGGTGAGGGTGGTGGGGTTAGAGTTGCTACTCGTCCAAC--TC-CAGTA
4090      4100      4110      4120      4130      4140      4150

   X
   A
   |
   AGGCCTACAATACCTGTGGAAACAGTAGGCCCCAGTGAAATTTTCCCCATA
    4160      4170      4180      4190      4200
```

Mon Dec 3 08:02:40 2001

```
> O <
O| |O  IntelliGenetics
> O <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file seq1.res made by sdavid on Wed 28 Nov 101 14:21:07-PST.

Query sequence being compared:                SEQ1 (1-22)
Number of sequences searched:                           6
Number of scores above cutoff:                          6

      Results of the initial comparison of SEQ1 (1-22) with:
        File : hpvcomplete.seq
```



PARAMETERS

Similarity matrix	Unitary	K-tuple	4
Mismatch penalty	1	Joining penalty	30
Gap penalty	1.00	Window size	8
Gap size penalty	0.33		
Cutoff score	1		
Randomization group	0		

SEARCH STATISTICS

	Mean	Median	Standard Deviation
Scores:	11	12	1.17
Times:	CPU 00:00:00.00		Total Elapsed 00:00:00.00

Number of residues:	47269
Number of sequences searched:	6
Number of scores above cutoff:	6

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
	**** 1 standard deviation above mean ****					
1. pph16	TOIG of: pph16 check: 6074	7904	13	13	1.71	0
	**** 0 standard deviation from mean ****					
2. a12360	TOIG of: a12360 check: 1580	7909	12	13	0.86	0
3. af131950	TOIG of: af131950 check: 557	7812	11	12	0.00	0
4. pph11	TOIG of: pph11 check: 3689	7931	11	12	0.00	0
5. hpu06714	TOIG of: hpu06714 check: 486	7801	10	11	-0.86	0
6. pph31a	TOIG of: pph31a check: 5866	7912	10	10	-0.86	0

1. SEQ1 (1-22)
   pph16 TOIG of: pph16 check: 6074 from: 1 to: 7904

TOIG of: pph16 check: 6074 from: 1 to: 7904

```
LOCUS       PPH16          7904 bp    DNA   circular  VRL      18-MAR-1994
DEFINITION  Human papillomavirus type 16 (HPV16), complete genome.
ACCESSION   K02718
VERSION     K02718.1  GI:333031
KEYWORDS    circular; complete genome.
SOURCE      Papilloma virus type 16 DNA, isolated from a human invasive
            cervical carcinoma.
  ORGANISM  Human papillomavirus type 16
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 7904)
  AUTHORS   Seedorf,K., Kraemmer,G., Duerst,M., Suhai,S. and Roewekamp,W.G.
  TITLE     Human papillomavirus type 16 DNA sequence
  JOURNAL   Virology 145, 181-185 (1985)
  MEDLINE   85246220
REFERENCE   2  (sites)
  AUTHORS   Kennedy,I.M., Haddow,J.K. and Clements,J.B.
  TITLE     A negative element in the human poapillomavirus type 16 genome acts
            at the level of late mRNA stability
  JOURNAL   J. Virol. 65, 2093-2097 (1991)
  MEDLINE   91162763
COMMENT     The sense strand of this double-stranded circular genome is shown,
            with a numbering system matching the first 60 bp of HPVa1, HPV6b
            and BPV1. The annotation of sites and features is solely based upon
            homology comparison with these other papillomaviruses. In addition
            to the coding sequences reported below, the authors note open
            reading frames which do not start with 'ATG', but which are found
            in other papillomaviruses. In particular, a second portion of the
            E1 gene may be located out to base 2813 (the E1 protein is thought
            to be generally involved in DNA replication).
            A potential 'CAT'-box region is found beginning at base 7895 below,
            and 'TATA' boxes for early and late transcripts may be located at
            17, 65 and 4289. Potential polyadenylation signals are at bases
            4213 and 7260.
            HPV16, in comparison to HPV types 6 and 11, is more often
            associated with malignant genital cancers in humans.
FEATURES             Location/Qualifiers
     source          1. .7904
                     /organism="Human papillomavirus type 16"
                     /db_xref="taxon:10581"
     TATA_signal     17. .23
     TATA_signal     65. .71
     gene            83. .559
                     /gene="E6"
     CDS             83. .559
                     /gene="E6"
                     /note="E6 ORF from 65 to 559; putative"
                     /codon_start=1
                     /product="transforming protein"
                     /protein_id="AAA46939.1"
                     /db_xref="GI:333032"
                     /translation="MHQKRTAMFQDPQERPRKLPQLCTELQTTIHDIILECVYCKQQL
                     LRREVYDFAFRDLCIVYRDGNPYAVCDKCLKFYSKISEYRHYCYSLYGTTLEQQYNKP
                     LCDLLIRCINCQKPLCPEEKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL"
     gene            562. .858
```

```
     CDS             /gene="E7"
                     562. .858
                     /gene="E7"
                     /note="E7 ORF from 544 to 858; putative"
                     /codon_start=1
                     /product="transforming protein"
                     /protein_id="AAA46940.1"
                     /db_xref="GI:333033"
                     /translation="MHGDTPTLHEYMLDLQPETTDLYCYEQLNDSSEEEDEIDGPAGQ
     gene            AEPDRAHYNIVTFCCKCDSTLRLCVQSTHVDIRTLEDLLMGTLGIVCPICSQKP"
     CDS             join(865. .1140,1140. .2813)
                     /gene="E1"
                     join(865. .1140,1140. .2813)
                     /gene="E1"
                     /note="E1 interrupted ORF from 859 to 2813; putative"
                     /codon_start=1
                     /product="replication protein"
                     /protein_id="AAA46936.1"
                     /db_xref="GI:459912"
                     /translation="MADPAGTNGEEGTGCNGWFYVEAVVEKKTGDAISDDENENDSDT
                     GEDLVDFIVNDNDYLTQAETETAHALFTAQEAKQHRDAVQVLKRKYLVSPLSDISGCV
                     DNNISPRLKAICIEKQSRAAKRRLFESEDSGYGNTEVETQQMLQVEGRHETETPCSQY
                     SGGSGGGCSQYSSGSGGEGVSERHTICQTPLTNILNVLKTSNAKAAMLAKFKELYGVS
                     FSELVRPPKSNKSTCCDWCIAAFGLTPSIADSIKTLLQQYCLYLHIQSLACSWGMVVL
                     LLVRYKCGKNRETIEKLLSKLLCVSPMCMMIEPPKLRSTAAALYWYKTGISNISEVYG
                     DTPEWIQRQTVLQHSFNDCTFELSQMVQWAYDNDIVDDSEIAYKYAQLADTNSNASAF
                     LKSNSQAKIVKDCATMCRHYKRAEKKQMSMSQWIKYRCDRVDDGGDWKQIVMFLRYQG
                     VEFMSFLTALKRFLQGIPKKNCILLYGAANTGKSLFGMSLMKFLQGSVICFVNSKSHF
                     WLQPLADAKIGMLDDATVPCWNYIDDNLRNALDGNLVSMDVKHRPLVQLKCPPLLITS
     gene            NINAGTDSRWPYLHNRLVVFTFPNEFPFDENGNPVYELNDKNWKSFFSRTWSRLSLHE
     CDS             DEDKENDGDSLPTFKCVSGQNTNTL"
                     2755. .3852
                     /gene="E2"
                     2755. .3852
                     /gene="E2"
                     /note="E2 ORF from 2725 to 3852; putative"
                     /codon_start=1
                     /product="regulatory protein"
                     /protein_id="AAA46941.1"
                     /db_xref="GI:333035"
                     /translation="METLCQRLNVCQDKILTHYENDSTDLRDHIDYWKHMRLECAIYY
                     KAREMGFKHINHQVVPTLAVSKNKALQAIELQLTLETIYNSQYSNEKWTLQDVSLEVY
                     LTAPTGCIKKHGYTVEVQFDGDICNTMHYTNWTHIYICEEASVTVVEGQVDYYGLYYV
                     HEGIRTYFVQFKDDAEKYSKNKVWEVHAGGQVILCPTSVFSSNEVSSPEIIRQHLANH
     gene            PAATHTKAVALGTEETQTTIQRPRSEPDTGNPCHTTKLLHRDSVDSAPILTAFNSSHK
     CDS             GRINCNSNTTPIVHLKGDANTLKCLRYRFKKHCTLYTAVSSTWHWTGHNVKHKSAIVT
                     LTYDSEWQRDQFLSQVKIPKTITVSTGFMSI"
                     3332. .3619
                     /gene="E4"
                     <3332. .3619
                     /gene="E4"
                     /note="E4 ORF from 3332 to 3619; putative"
                     /codon_start=1
                     /protein_id="AAA46937.1"
                     /db_xref="GI:459913"
                     /translation="YYVLHLCLAATKYPLLKLLGSTWPTTPPRPIPKPSPWAPKKHRR
     gene            LSSDQDQSQTPETPATPLSCCTETQWTVLQSSLHLTAHTKDGLTVIVTLHP"
     CDS             3863. .4099
                     /gene="E5"
                     <3863. .4099
                     /gene="E5"
                     /note="E5 ORF from 3863 to 4099; putative"
                     /codon_start=1
                     /protein_id="AAA46938.1"
                     /db_xref="GI:459914"
                     /translation="YCIHNITGVLFALLCVLLCVCLLIRPLLLSVSTYTSLIILVLLL
     polyA_signal    WITAASAFRCFIVYIIFVYIPLFLIHTHARFLIT"
                     4213. .4218
     gene            /note="putative"
                     4235. .5656
     CDS             /gene="L2"
                     4235. .5656
                     /gene="L2"
                     /note="L2 ORF from 4133 to 5656; putative"
                     /codon_start=1
                     /product="minor capsid protein"
                     /protein_id="AAA46942.1"
                     /db_xref="GI:333036"
                     /translation="MRHKRSAKRTKRASATQLYKTCKQAGTCPPDIIPKVEGKTIAEQ
                     ILQYGSMGVFFGGLGIGTGSGTGGRTGYIPLGTRPPTATDTLAPVRPPLTVDPVGPSD
                     PSIVSLVEETSFIDAGAPTSVPSIPPDVSGFSITTSTDTTPAILDINNTVTTVTTHNN
                     PTFTDPSVLQPPTPAETGGHFTLSSSTISTHNYEEIPMDTFIVSTNPNTVTSSTPIPG
                     SRPVARLGLYSRTTQQVKVVDPAFVTTPTKLITYDNPAYEGIDVDNTLYFSSNDNSIN
                     IAPDPDFLDIVALHRPALTSRRTGIRYSRIGNKQTLRTRSGKSIGAKVHYYYDLSTID
                     PAEEIELQTITPSTYTTTSHAASPTSINNGLYDIYADDFITDTSTTPVPSVPSTSLSG
                     YIPANTTIPFGGAYNIPLVSGPDIPINITDQAPSLIPIVPGSPQYTIIADAGDFYLHP
                     SYYMLRKRRKRLPYFFSDVSLAA"
     TATA_signal     4289. .4295
     gene            /gene="L2"
                     5559. .7154
     CDS             /gene="L1"
                     5559. .7154
                     /gene="L1"
                     /note="L1 ORF from 5526 to 7154; putative"
                     /codon_start=1
                     /product="major capsid protein"
                     /protein_id="AAA46943.1"
                     /db_xref="GI:333037"
                     /translation="MQVTFIYILVITCYENDVNVYHIFFQMSLWLPSEATVYLPPVPV
                     SKVVSTDEYVARTNIYYHAGTSRLLAVGHPYFPIKKPNNNKILVPKVSGLQYRVFRIH
                     LPDPNKFGFPDTSFYNPDTQRLVWACVGVEVGRGQPLGVGISGHPLLNKLDDTENASA
                     YAANAGVDNRECISMDYKQTQLCLIGCKPPIGEHWGKGSPCTNVAVNPGDCPPLELIN
                     TVIQDGDMVHTGFGAMDFTTLQANKSEVPLDICTSICKYPDYIKMVSEPYGDSLFFYL
                     RREQMFVRHLFNRAGTVGENVPDDLYIKGSGSTANLASSNYFPTPSGSMVTSDAQIFN
                     KPYWLQRAQGHNNGICWGNQLFVTVVDTTRSTNMSLCAAISTSETTYKNTNFKEYLRH
                     GEEYDLQFIFQLCKITLTADVMTYIHSMNSTILEDWNFGLQPPPGGTLEDTYRFVTQA
                     IACQKHTPPAPKEDDPLKKYTFWEVNLKEKFSADLDQFPLGRKFLLQAGLKAKPKFTL
                     GKRKATPTTSSTSTTAKRKKRKL"
     polyA_signal    7260. .7265
BASE COUNT      2601 a    1377 c    1509 g    2417 t
ORIGIN      Unreported.

  PPH16  Length: 7904  November 28, 2001 14:10  Type: N  Check: 6074  ..

Initial Score   =     13  Optimized Score  =     13  Significance  =  1.71
Residue Identity =    59%  Matches          =     13  Mismatches    =     9
Gaps            =      0  Conservative Substitutions               =     0

                                                          X       10        20 X
                                                          TGACTGTGAACGTTCGAGATGA
  TCAAGTTGACTATTATGGTTTATATTATGTTCATGAAGGAATACGAACATATTTTGTGCAGTTTAAAGATGA
     3210      3220      3230      3240      3250    X  3260      3270    X

  TGCAGAAAAATATAGTAAAAATAAAGTATGGGAAGTTCATGCGGGTGGTC
     3280      3290      3300      3310      3320

2. SEQ1 (1-22)
   a12360       TOIG of: a12360  check: 1580  from: 1  to: 7909

  TOIG of: a12360  check: 1580  from: 1  to: 7909

LOCUS       A12360          7909 bp    DNA                     PAT       12-DEC-1993
DEFINITION  complete nucleotide sequence of HPV-33.
ACCESSION   A12360
VERSION     A12360.1  GI:492936
KEYWORDS    .
SOURCE      Human papillomavirus type 33.
  ORGANISM  Human papillomavirus type 33
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 7909)
  AUTHORS   .
  JOURNAL   Patent: WO 8705630-A 1 24-SEP-1987;
FEATURES             Location/Qualifiers
```

```
     source          1. .7909
                     /organism="Human papillomavirus type 33"
                     /db_xref="taxon:10586"
BASE COUNT      2544 a    1354 c    1535 g    2474 t      2 others
ORIGIN

A12360  Length: 7909  November 28, 2001 14:10  Type: N  Check: 1580  ..

Initial Score    =     12  Optimized Score  =     13  Significance =  0.86
Residue Identity =    59%  Matches          =     13  Mismatches   =     9
Gaps             =      0  Conservative Substitutions               =     0

                                                    X       10        20 X
                                                    TGACTGTGAACGTTCGAGATGA
                                                    | |||||| |   |    ||||
   GGTGTGGCTTTGTGAACCACCAAAATGTTTTAAAAAACAAGGAGAAACAGTAACTGTGCAATATGACAATGA
 3050      3060      3070      3080      3090      3100      3110        X

   CAAAAAAAATACAATGGATTATACAAACTGGGGTGAAATATATATTATAG
 3120      3130      3140      3150      3160


3. SEQ1 (1-22)
   af131950     TOIG of: af131950  check: 557  from: 1  to: 7812

  TOIG of: af131950  check: 557  from: 1  to: 7812

  LOCUS       AF131950     7812 bp    DNA             VRL       07-FEB-2001
  DEFINITION  Human papillomavirus candHPV85, complete genome.
  ACCESSION   AF131950
  VERSION     AF131950.1  GI:4574720
  KEYWORDS    .
  SOURCE      Human papillomavirus candHPV85.
    ORGANISM  Human papillomavirus candHPV85
              Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
              Papillomavirus.
  REFERENCE   1  (bases 1 to 7812)
    AUTHORS   Chow,V.T.K. and Leong,P.W.F.
    TITLE     Complete nucleotide sequence, genomic organization and phylogenetic
              analysis of a novel genital human papillomavirus type, HLT7474-S
    JOURNAL   J. Gen. Virol. 80 (Pt 11), 2923-2929 (1999)
    MEDLINE   20047972
     PUBMED   10580054
  REFERENCE   2  (bases 1 to 7812)
    AUTHORS   Chow,V.T.K. and Leong,W.F.
    TITLE     Direct Submission
    JOURNAL   Submitted (26-FEB-1999) Department of Microbiology, National
              University of Singapore, 10 Kent Ridge Crescent, Singapore 119260,
              Singapore
  FEATURES             Location/Qualifiers
       source          1. .7812
                       /organism="Human papillomavirus candHPV85"
                       /db_xref="taxon:151757"
                       /note="isolated from scraped uterine cervical cells from a
                       female sex worker; overlapping PCR products"
       protein_bind    40. .51
                       /note="putative"
                       /bound_moiety="E2 protein"
                       /function="transcriptional regulation"
       protein_bind    56. .67
                       /note="putative"
                       /bound_moiety="E2 protein"
                       /function="transcriptional regulation"
       TATA_signal     71. .79
                       /note="putative"
       gene            105. .578
                       /gene="E6"
       CDS             105. .578
                       /gene="E6"
                       /codon_start=1
                       /product="putative transforming protein E6"
                       /protein_id="AAD24181.1"
                       /db_xref="GI:4574721"
                       /translation="MAEFGNPATRPYKLPDLCNTLDTSLQDIEISCVYCKSVLQRTEV
                       YEFAFADLFVVYRDGIPYAACQNCLMFYSKIRELRYYSDSVYGETLEKLTNSNIYDIL
                       IRCLRCQKPLCPAEKLKHLNEKRRFHKIAGKYRGQCRRCMTRAQEQQGSRRETQV"
       gene            587. .913
                       /gene="E7"
       CDS             587. .913
                       /gene="E7"
                       /codon_start=1
                       /product="putative transforming protein E7"
                       /protein_id="AAD24182.1"
                       /db_xref="GI:4574722"
                       /translation="MHGPKPTVHDIVLDLEPYNEVQEVDLYCYEELNNSEEEIDEPDN
                       AINHRQPLLARREELQRHTICCVCCKCEASLQLVVESSAADLRDLQQLFLGTLSFLCP
                       LCAVLR"
       gene            920. .2866
                       /gene="E1"
       CDS             920. .2866
                       /gene="E1"
                       /codon_start=1
                       /product="putative replication protein E1"
                       /protein_id="AAD24183.1"
                       /db_xref="GI:4574723"
                       /translation="MADPEGTDGDGTGCNGWFFVQAIVDKKTGDTISEDEDEDGTDTG
                       SDLVDFIDTSNTNYMQADREAAQALLHAQEVETDTKLLHALKRKYGAHSTENSPCRDT
                       ASIHSNLSSPLQEISLNSYNNTAKRRLCSVPDSGYGNTQVETLQTQVTLDTNVFGDGK
                       NGDGLNSEACSTDNEMDIENQNPNSPMTQIVSLLKVNNKKAAILAKFKETYGLSFTDL
                       VRTFKSDKTTCTDWVAAICGVNPNIAEGFKTLIQPYVLYAHIQCMDCSWGVFILALLR
                       YKCGKNRLTVAKGLSTLLHVPDTHMLIEPPKLRSSCAALYWYRTGISNISEVTGDTPE
                       WIQRQTIIQHGIDDSVFDLSEMIQWAFDNDYIDESDIAYEYAQLADCNSNAAAFLKSN
                       CQAKYLRDCAVMCRHYKRAQRKQMNMSQWISYRCDKIDDGGDWKPIVQFLRFQGIEFI
                       TFLRAFKDFLKGTPKKNCIVIYGPANTGKSYFCMSLIQFLHGTVLSFVNSNSHFWLEP
                       LTDTKIAMVDDATPTCWSYFDNYMRNALDGNPISIDRKHKHLIQMKCPPMLITSNTNP
                       ATDDRWPYLRSRVTVFTFPHTFPFDSNGNPVYDINDKNWKCFFKRTWSRLDLHQEEED
                       TENDGCPIPTFKCVTGKDSRTLRD"
       gene            2799. .3932
                       /gene="E2"
       CDS             2799. .3932
                       /gene="E2"
                       /codon_start=1
                       /product="putative regulatory protein E2"
                       /protein_id="AAD24184.1"
                       /db_xref="GI:4574724"
                       /translation="MMDALSQRLNVLQEKILEHYETDSTDICDQIDYWKCVRLENAIY
                       YAAREHGLKTINHHVVPTFQISKSKAHEAIELQMALESLAKSEFKNELWTLQDTCQEL
                       YQTPPQQCFKKQGQTVEVRYDGDKDNTMHYTSWDYIYYVTEGDKWCKTKGYVNYCGLY
                       YIKEGQQTYYVQFKCDAQQYGQSGKWEVWYNGKKIECVNTCDKIIQCSESMYSTCDET
                       VSATAIARELQHPTTPYTEATTVCTQKSGGSAPTRNPFRHCGFTETSEVDGLSVDHLN
                       NPLLSAPAGNNGFRKHNSGDTTPIVHLKGDKNRLKCLRYRLQKFHNLYTQISCTWHWI
                       GGKETTRTGILTLAYANETQRQKFLDVVKIPNTVNVSLGYMTL"
       gene            <3412. .3699
                       /gene="E4"
       CDS             <3412. .3699
                       /gene="E4"
                       /note="start codon absent"
                       /codon_start=1
                       /product="putative protein E4"
                       /protein_id="AAD24185.1"
                       /db_xref="GI:4574725"
                       /translation="YNVLNLCTVPVTRQYPLLQLLENYNTPPHRIPKPPPCAPKRAGV
                       RRRLETHSDIVDSQKQVKSTDCPWTTSTTPCLVHLQATTGSGSTIVVTLRL"
       gene            3940. .4161
                       /gene="E5"
       CDS             3940. .4161
                       /gene="E5"
                       /codon_start=1
                       /product="putative protein E5"
                       /protein_id="AAD24186.1"
                       /db_xref="GI:4574726"
                       /translation="MLIAVFLLCFCVCLYVCCTIPLFPATNVCAHAWLLVFLLIVFIT
                       TPFTAFIVYIFFFILPMFLLHLHSVHTFD"
       gene            4226. .5629
                       /gene="L2"
       CDS             4226. .5629
```

```
                     /gene="L2"
                     /codon_start=1
                     /product="putative minor capsid protein L2"
                     /protein_id="AAD24187.1"
                     /db_xref="GI:4574727"
                     /translation="MVSHRAARRKRASATDLYKTCKQSGTCPPDVINKVEGTTLADKL
                     LQWTSLGIFLGGLGIGTGSGTGGRTGYIPLGGRPNTVVDVSPARPPVVIESVGPSDPS
                     IVTLVEESSIVTSGAPVPTFTGTSGFEITSSATTTPAVLDITPASGSVQLSSTSFTNP
                     AFTDPSVIEVPQTGEVSGDIFITTPTSGTHGYEEIPMHTFATQGRGTEPISSTPIPGV
                     RRVAGPRLYSQAYQQVKITNSDFISRPSTLVTFTNPAYEPIDTTLTFSPQDVVPDPDF
                     MDIVRLHRPALTSRRGTVRFSRLGKKLTMSTRSGKQIGAQVHYYHDISPISHIGESIE
                     MQPLLPDAAVTADTNGLFDIYADTDIDNNAMLYDRNISDVTQPTTSTISSVSSRYSNT
                     TIPLATSWDVPVHTGPDMTLPTTIPQWPNIVPLLPNNTHSVVLQGTNYYLWPNYYFIF
                     KKRKRVPYFLTDGFVAF"
     gene            5610. .7124
                     /gene="L1"
     CDS             5610. .7124
                     /gene="L1"
                     /codon_start=1
                     /product="putative major capsid protein L1"
                     /protein_id="AAD24188.1"
                     /db_xref="GI:4574728"
                     /translation="MALWRSSDSKVYLPPPSVAKVVNTDDYVTRTSTFYHAGSSRLLT
                     VGHPYYKVTSNGGRKQDIPKVSAYQYRVFRVTLPDPNKFGLPDTNIYNPETQRLVWAC
                     VGMEVGRGQPLGVSLSGHPFYNKLDDTENSHVATSVVTHDTRDNVSVDYKQTQLCIIG
                     CVPAIGEHWAKGTACKPGAVQTGDCPPLELVNTPIEDGDMIDTGYGAMDFSTLQDNKS
                     EVPLDICQSICKYPDYLQMSADAYGDSMFFCLRREQLFARHFWNRGGTIGDAVPETLY
                     IKGTNDRATPGSCIYSPSPSGSMVSSDAQMFNKPYWLHKAQGQNNGICWHNQLFITVV
                     DTTRSTNLTLSTATTNPVPSIYEPSKFKEYTRHVEEYDLQFIFQLCKITLTTDVMSYI
                     HNMDPTILDSWNFGVSPPPSASLVDTYRFLQSSAITCQKDVVVPQKKDPYEKLKFWNV
                     DLKEHFSSDLDQFPLGRKFLLQAGLRPKPTIGPRKRVASTSTATRPSKRKRTAK"
     protein_bind    7456. .7467
                     /note="putative"
                     /bound_moiety="E2 protein"
                     /function="transcriptional regulation"
     misc_feature    7647. .7666
                     /note="consensus motif of oncogenic anogenital HPV types"
     protein_bind    7776. .7787
                     /note="putative"
                     /bound_moiety="E2 protein"
                     /function="transcriptional regulation"
BASE COUNT      2502 a    1391 c    1558 g    2361 t
ORIGIN

AF131950  Length: 7812  November 28, 2001 14:10  Type: N  Check: 557  ..

Initial Score   =     11  Optimized Score  =     12  Significance =   0.00
Residue Identity =    56%  Matches         =     13  Mismatches   =      9
Gaps            =      1  Conservative Substitutions               =      0


                                                    X       10        20
                                                    TGACTGTGAACGTTCGAG-ATG
   AATACAATGTTCTGAATCTATGTACAGTACCTGTGACGAGACAGTATCCGCTACTGCAATTGCTAGAGAATT
3410      3420      3430      3440      3450      3460      3470      3480

   X
   A
   |
   ACAACACCCCACCACACCGTATACCGAAGCCACCACCGTGTGCACCCAAAA
   X    3490      3500      3510      3520      3530


4. SEQ1 (1-22)
   pph11        TOIG of: pph11  check: 3689  from: 1  to: 7931

TOIG of: pph11  check: 3689  from: 1  to: 7931

LOCUS       PPH11          7931 bp    DNA   circular  VRL       02-JUN-1994
DEFINITION  Human papillomavirus type 11 (HPV-11) complete genome.
ACCESSION   M14119
VERSION     M14119.1  GI:333026
KEYWORDS    complete genome.
```

```
SOURCE      Human laryngeal papillomavirus type 11 DNA.
  ORGANISM  Human papillomavirus type 11
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 7931)
  AUTHORS   Dartmann,K., Schwarz,E., Gissmann,L. and Zur Hausen,H.
  TITLE     The nucleotide sequence and genome organization of human papilloma
            virus type 11
  JOURNAL   Virology 151, 124-130 (1986)
  MEDLINE   86181601
COMMENT     ORF L1 is assumed to encode the major structural protein.
FEATURES             Location/Qualifiers
     source          1. .7931
                     /organism="Human papillomavirus type 11"
                     /db_xref="taxon:10580"
     CAAT_signal     9. .15
     protein_bind    35. .46
                     /note="putative"
                     /function="gene transcription"
                     /bound_moiety="E2"
     protein_bind    50. .61
                     /note="putative"
                     /function="gene transcription"
                     /bound_moiety="E2"
     TATA_signal     66. .71
                     /note="putative"
     gene            102. .554
                     /gene="E6"
     CDS             102. .554
                     /gene="E6"
                     /note="102 is position of first start codon in ORF E6;
                     putative"
                     /codon_start=1
                     /product="transforming protein"
                     /protein_id="AAA46927.1"
                     /db_xref="GI:496193"
                     /translation="MESKDASTSATSIDQLCKTFNLSLHTLQIQCVFCRNALTTAEIY
                     AYAYKNLKVVWRDNFPFAACACCLELQGKINQYRHFNYAAYAPTVEEETNEDILKVLI
                     RCYLCHKPLCEIEKLKHILGKARFIKLNNQWKGRCLHCWTTCMEDLLP"
     gene            530. .826
                     /gene="E7"
     CDS             530. .826
                     /gene="E7"
                     /note="530 is position of first start codon in ORF E7;
                     putative"
                     /codon_start=1
                     /product="transforming protein"
                     /protein_id="AAA46928.1"
                     /db_xref="GI:496194"
                     /translation="MHGRLVTLKDIVLDLQPPDPVGLHCYEQLEDSSEDEVDKVDKQD
                     AQPLTQHYQILTCCCGCDSNVRLVVECTDGDIRQLQDLLLGTLNIVCPICAPKP"
     gene            832. .2781
                     /gene="E1"
     CDS             832. .2781
                     /gene="E1"
                     /note="832 is position of first start codon in ORF E1;
                     putative"
                     /codon_start=1
                     /product="replication protein"
                     /protein_id="AAA46929.1"
                     /db_xref="GI:496195"
                     /translation="MADDSGTENEGSGCTGWFMVEAIVEHTTGTQISEDEEEEVEDSG
                     YDMVDFIDDRHITQNSVEAQALFNRQEADAHYATVQDLKRKYLGSPYVSPISNVANAV
                     ESEISPRLDAIKLTTQPKKVKRRLFETRELTDSGYGYSEVEAATQVEKHGDPENGGDG
                     QERDTGRDIEGEGVEHREAEAVDDSTREHADTSGILELLKCKDIRSTLHGKFKDCFGL
                     SFVDLIRPFKSDRTTCADWVVAGFGIHHSIADAFQKLIEPLSLYAHIQWLTNAWGMVL
                     LVLIRFKVNKSRCTVARTLGTLLNIPENHMLIEPPKIQSGVRALYWFRTGISNASTVI
                     GEAPEWITRQTVIEHSLADSQFKLTEMVQWAYDNDICEESEIAFEYAQRGDFDSNARA
                     FLNSNMQAKYVKDCAIMCRHYKHAEMKKMSIKQWIKYRGTKVDSVGNWKPIVQFLRHQ
                     NIEFIPFLSKLKLWLHGTPKKNCIAIVGPPDTGKSCFCMSLIKFLGGTVISYVNSCSH
                     FWLQPLTDAKVALLDDATQPCWTYMDTYMRNLLDGNPMSIDRKHRALTLIKCPPLLVT
                     SNIDISKEEKYKYLHSRVTTFTFPNPFPFDRNGNAVYELSDANWKCFFERLSSSLDIE
```

```
                DSEDEEDGSNSQAFRCVPGSVVRTL"
gene            2723. .3826
                /gene="E2"
CDS             2723. .3826
                /gene="E2"
                /note="2723 is position of first start codon in ORF E2;
                putative"
                /codon_start=1
                /product="regulatory protein"
                /protein_id="AAA46930.1"
                /db_xref="GI:496196"
                /translation="MEAIAKRLDACQDQLLELYEENSIDIHKHIMHWKCIRLESVLLH
                KAKQMGLSHIGLQVVPPLTVSETKGHNAIEMQMHLESLAKTQYGVEPWTLQDTSYEMW
                LTPPKRCFKKQGNTVEVKFDGCEDNVMEYVVWTHIYLQDNDSWVKVTSSVDAKGIYYT
                CGQFKTYYVNFNKEAQKYGSTNHWEVCYGSTVICSPASVSSTVREVSIAEPTTYTPAQ
                TTAPTVSACTTEDGVSAPPRKRARGPSTNNTLCVANIRSVDSTINNIVTDNYNKHQRR
                NNCHSAATPIVQLQGDSNCLKCFRYRLNDKYKHLFELASSTWHWASPEAPHKNAIVTL
                TYSSEEQRQQFLNSVKIPPTIRHKVGFMSLHLL"
gene            3255. .3581
                /gene="E4"
CDS             3255. .3581
                /gene="E4"
                /note="3255 is position of first start codon in ORF E4;
                putative"
                /codon_start=1
                /protein_id="AAA46931.1"
                /db_xref="GI:496197"
                /translation="MVVPIIGKYVMAAQLYVLLHLYLALYEKYPLLNLLHTPPHRPPP
                LQCPPAPRKTACRRRLGSEHVDRPLTTPCVWPTSDPWTVQSTTSSLTITTSTKEGTTV
                TVQLRL"
gene            3871. .4146
                /gene="E5A"
CDS             3871. .4146
                /gene="E5A"
                /note="3871 is position of first start codon in ORFE5A;
                putative"
                /codon_start=1
                /protein_id="AAA46932.1"
                /db_xref="GI:496198"
                /translation="MEVVPVQIAAATTTTLILPVVIAFAVCILSIVLIILISDFVVYT
                SVLVLTLLLYLLLWLLLTTPLQFFLLTLCVCYFPAFYIHIYIVQTQQ"
gene            4146. .4370
                /gene="E5B"
CDS             4146. .4370
                /gene="E5B"
                /note="4146 is position of first start codon in ORF E5B;
                putative"
                /codon_start=1
                /protein_id="AAA46933.1"
                /db_xref="GI:496199"
                /translation="MVMLTCHLNDGDTWLFLWLFTAFVVAVLGLLLLHYRAVHGTEKT
                KCAKCKSNRNTTVDYVYMSHGDNGDYVYMN"
polyA_signal    4371. .4376
                /note="putative"
gene            4417. .5784
                /gene="L2"
CDS             4417. .5784
                /gene="L2"
                /note="4417 is position of first start codon in ORF L2;
                putative"
                /codon_start=1
                /product="minor capsid protein"
                /protein_id="AAA46934.1"
                /db_xref="GI:496200"
                /translation="MKPRARRRKRASATQLYQTCKATGTCPPDVIPKVEHTTIADQIL
                KWGSLGVFFGGLGIGTGAGSGGRAGYIPLGSSPKPAITGGPAARPPVLVEPVAPSDPS
                IVSLIEESAIINAGAPEVVPPTQGGFTITSSESTTPAILDVSVTNHTTTSVFQNPLFT
                EPSVIQPQPPVEASGHILISAPTITSQHVEDIPLDTFVVSSSDSGPTSSTPLPRAFPR
                PRVGLYSRALQQVQVTDPAFLSTPQRLVTYDNPVYEGEDVSLQFTHESIHNAPDEAFM
                DIIRLHRPAITSRRGLVRFSRIGQRGSMYTRSGQHIGARIHYFQDISPVTQAAEEIEL
                HPLVAAENDTFDIYAEPFDPIPDPVQHSVTQSYLTSTPNTLSQSWGNTTVPLSIPSDW
                FVQSGPDITFPTASMGTPFSPVTPALPTGPVFITGSDFYLHPTWYFARRRRKRIPLFF
                TDVAA"
polyA_signal    4545. .4551
                /gene="L2"
                /note="putative"
gene            5771. .7276
                /gene="L1"
CDS             5771. .7276
                /gene="L1"
                /note="5771 is position of first start codon in ORF L1;
                putative"
                /codon_start=1
                /product="major capsid protein"
                /protein_id="AAA46935.1"
                /db_xref="GI:496201"
                /translation="MWRPSDSTVYVPPPNPVSKVVATDAYVKRTNIFYHASSSRLLAV
                GHPYYSIKKVNKTVVPKVSGYQYRVFKVVLPDPNKFALPDSSLFDPTTQRLVWACTGL
                EVGRGQPLGVGVSGHPLLNKYDDVENSGGYGGNPGQDNRVNVGMDYKQTQLCMVGCAP
                PLGEHWGKGTQCSNTSVQNGDCPPLELITSVIQDGDMVDTGFGAMNFADLQTNKSDVP
                LDICGTVCKYPDYLQMAADPYGDRLFFYLRKEQMFARHFFNRAGTVGEPVPDDLLVKG
                GNNRSSVASSIYVHTPSGSLVSSEAQLFNKPYWLQKAQGHNNGICWGNHLFVTVVDTT
                RSTNMTLCASVSKSATYTNSDYKEYMRHVEEFDLQFIFQLCSITLSAEVMAYIHTMNP
                SVLEDWNFGLSPPPNGTLEDTYRYVQSQAITCQKPTPEKEKQDPYKDMSFWEVNLKEK
                FSSELDQFPLGRKFLLQSGYRGRTSARTGIKRPAVSKPSTAPKRKRTKTKK"
5'UTR           join(7277. .7931,1. .101)
                /standard_name="LCR"
                /function="regulatory region"
repeat_region   7320. .7360
                /standard_name="direct repeats"
                /note="putative"
repeat_region   7339. .7374
                /standard_name="direct repeats"
                /note="putative"
repeat_region   7374. .7403
                /note="putative"
polyA_signal    7457. .7462
                /note="putative"
protein_bind    7592. .7603
                /note="putative"
                /function="gene transcription"
                /bound_moiety="E2"
polyA_signal    7748. .7753
                /note="putative"
protein_bind    7890. .7901
                /note="putative"
                /function="gene transcription"
                /bound_moiety="E2"
BASE COUNT      2406 a    1519 c    1736 g    2270 t
ORIGIN      4557 bp upstream of HindIII site.

  PPH11  Length: 7931  November 28, 2001 14:10  Type: N  Check: 3689  ..

Initial Score     =     11  Optimized Score  =     12  Significance  =  0.00
Residue Identity  =     56%  Matches         =     13  Mismatches    =     9
Gaps              =      1  Conservative Substitutions               =     0

                                                          X       10        20
                                                          TG-ACTGTGAACGTTCGAGATG
                                                          || ||||   ||   |||||
    TTTAATCTTTCTTTGCACACTCTGCAAATTCAGTGCGTGTTTTGCAGGAATGCACTGACCACCGCAGAGATA
  160       170       180       190       200       210       220       230

    X
    A

    TATGCATATGCCTATAAGAACCTAAAGGTTGTGTGGCGAGACAACTTTCCC
    X       240       250       260       270       280


5. SEQ1 (1-22)
   hpu06714     TOIG of: hpu06714  check: 4862  from: 1  to: 7801

  TOIG of: hpu06714  check: 4862  from: 1  to: 7801

LOCUS       HPU06714      7801 bp     DNA              VRL       04-FEB-1997
```

```
DEFINITION  Human papillomavirus HPV-1A (3-3), complete genome.
ACCESSION   U06714
VERSION     U06714.1  GI:458704
KEYWORDS    .
SOURCE      Human papillomavirus.
  ORGANISM  Human papillomavirus
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 7801)
  AUTHORS   Danos,O., Katinka,M. and Yaniv,M.
  TITLE     Human papillomavirus 1A complete DNA sequence: a novel type of
            genome organization among Papovaviridae
  JOURNAL   EMBO J. 1, 231-236 (1982)
  MEDLINE   84182467
REFERENCE   2  (bases 1 to 7801)
  AUTHORS   Meissner,J.
  TITLE     Complete nucleotide sequencing of an HPV-1a variant and
            determination of extant errors in the prototype HPV-1a sequence
  JOURNAL   Virus Genes 9 (2), 189-191 (1995)
  MEDLINE   95250312
REFERENCE   3  (bases 1 to 7801)
  AUTHORS   Meissner,J.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-FEB-1994) John D. Meissner, Duke University Medical
            Center, Microbiology, 277 Carl Building, Research Drive, Durham, NC
            27710 USA
FEATURES             Location/Qualifiers
     source          1. .7801
                     /organism="Human papillomavirus"
                     /strain="HPV-1A (3-3)"
                     /db_xref="taxon:10566"
     variation       142
                     /replace="a"
     conflict        1283
                     /citation=[1]
                     /replace="a"
     conflict        2301
                     /citation=[1]
                     /replace="t"
     variation       2825
                     /replace="t"
     conflict        3884. .3886
                     /citation=[1]
                     /replace="aa"
     conflict        4332
                     /citation=[1]
                     /replace="a"
     conflict        4376. .4382
                     /citation=[1]
                     /replace="gagggaa"
     variation       5794
                     /replace="a"
     variation       6905
                     /replace="g"
     variation       7186. .7187
                     /note="15 bp deletion"
                     /citation=[1]
                     /replace="ctagattgtattgctat"
     variation       7560
                     /replace="c"
     variation       7618
                     /replace="g"
     conflict        7677. .7678
                     /citation=[1]
                     /replace="cc"
     variation       7787
                     /replace="g"
BASE COUNT      2389 a   1482 c   1664 g   2266 t
ORIGIN

HPU06714  Length: 7801  November 28, 2001 14:10  Type: N  Check: 4862  ..

Initial Score  =      10  Optimized Score  =      11  Significance = -0.86

Residue Identity =    55%  Matches          =      15  Mismatches  =      6
Gaps             =      6  Conservative Substitutions              =      0

                                                          X        10
                                                          TGACT--GTGAACGTTCGAG--
                                                          ||||   || | |  | ||
    TAAGTGACCAAGAAGACGAGGGAAACGATGGAAAACCTCAGCAGTCGCTTAGACTTACTGCAAGAGC-AGCT
      2570      2580      2590      2600      2610    X 2620     2630

     20 X
    -ATGA
     ||||
    AATGAACCTATATGAACAGGACAGTAAATTGATAGAAGATCAAATTAAGCAGTGG
     2640      2650      2660      2670      2680      2690

6. SEQ1 (1-22)
   pph31a       TOIG of: pph31a  check: 5866  from: 1  to: 7912

  TOIG of: pph31a  check: 5866  from: 1  to: 7912

LOCUS       PPH31A       7912 bp    DNA   circular  VRL       18-MAR-1994
DEFINITION  Human papillomavirus type 31 (HPV-31) complete genome.
ACCESSION   J04353
VERSION     J04353.1  GI:333048
KEYWORDS    complete genome.
SOURCE      Human papillomavirus type 31 DNA.
  ORGANISM  Human papillomavirus type 31
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 7912)
  AUTHORS   Goldsborough,M.D., DiSelvestre,D., Temple,G.F. and Lorincz,A.T.
  TITLE     Nucleotide sequence of human papillomavirus type 31: A cervical
            neoplasia associated virus
  JOURNAL   Virology 171, 306-311 (1989)
  MEDLINE   89299478
COMMENT     Draft entry and computer-readable copy of sequence [1] kindly
            submitted by M.D.Goldsborough, 05-JUL-1989.
FEATURES             Location/Qualifiers
     source          1. .7912
                     /organism="Human papillomavirus type 31"
                     /db_xref="taxon:10585"
     TATA_signal     19. .24
     TATA_signal     69. .74
     gene            108. .557
                     /gene="E6"
     CDS             108. .557
                     /gene="E6"
                     /note="ORF E6 from bp 39 to 557"
                     /codon_start=1
                     /product="transforming protein"
                     /protein_id="AAA46950.1"
                     /db_xref="GI:459916"
                     /translation="MFKNPAERPRKLHELSSALEIPYDELRLNCVYCKGQLTETEVLD
                     FAFTDLTIVYRDDTPHGVCTKCLRFYSKVSEFRWYRYSVYGTTLEKLTNKGICDLLIR
                     CITCQRPLCPEEKQRHLDKKKRFHNIGGRWTGRCIACWRRPRTETQV"
     misc_feature    228. .236
                     /gene="E6"
                     /standard_name="Splice donor"
     misc_feature    403. .414
                     /gene="E6"
                     /standard_name="Splice acceptor"
     gene            560. .856
                     /gene="E7"
     CDS             560. .856
                     /gene="E7"
                     /note="ORF E7 from bp 545 to 856"
                     /codon_start=1
                     /product="transforming protein"
                     /protein_id="AAA46951.1"
                     /db_xref="GI:459917"
                     /translation="MRGETPTLQDYVLDLQPEATDLHCYEQLPDSSDEEDVIDSPAGQ
                     AEPDTSNYNIVTFCCQCKSTLRLCVQSTQVDIRILQELLMGSFGIVCPNCSTRL"
```

```
gene            862. .2751
                /gene="E1"
CDS             862. .2751
                /gene="E1"
                /note="ORF E1 from bp 850 to bp 2751"
                /codon_start=1
                /product="replication protein"
                /protein_id="AAA46952.1"
                /db_xref="GI:459918"
                /translation="MADPAGTDGEGTGCNGWFYVEAVIDRQTGDNISEDENEDSSDTG
                EDMVDFIDNCNVYNNQAEAETAQALFHAQEAEEHAEAVQVLKRKYVGSPLSDISSCVD
                YNISPRLKAICIENNSKTAKRRLFELPDSGYGNTEVETQQMVQVEEQQTTLSCNGSDG
                THSERENETPTRNILQVLKTSNGKAAMLGKFKELYGVSFMELIRPFQSNKSTCTDWCV
                AAFGVTGTVAEGFKTLLQPYCLYCHLQSLACSWGMVMLMLVRFKCAKNRITIEKLLEK
                LLCISTNCMLIQPPKLRSTAAALYWYRTGMSNISDVYGETPEWIERQTVLQHSFNDTT
                FDLSQMVQWAYDNDVMDDSEIAYKYAQLADSDSNACAFLKSNSQAKIVKDCGTMCRHY
                KRAEKRQMSMGQWIKSRCDKVSDEGDWRDIVKFLRYQQIEFVSFLSALKLFLKGVPKK
                NCILIHGAPNTGKSYFGMSLISFLQGCIISYANSKSHFWLQPLADAKIGMLDDATTPC
                WHYIDNYLRNALDGNPVSIDVKHKALMQLKCPPLLITSNINAGKDDRWPYLHSRLVVF
                TFPNPFPFDKNGNPVYELSDKNWKSFFSRTWCRLNLHEEEDKENDGDSFSTFKCVSGQ
                NIRTL"
gene            2693. .3811
                /gene="E2"
CDS             2693. .3811
                /gene="E2"
                /note="ORF E2 from bp 2663 to 3811"
                /codon_start=1
                /product="regulatory protein"
                /protein_id="AAA46953.1"
                /db_xref="GI:459919"
                /translation="METLSQRLNVCQDKILEHYENDSKRLCDHIDYWKHIRLECVLMY
                KAREMGIHSINHQVVPALSVSKAKALQAIELQMMLETLNNTEYKNEDWTMQQTSLELY
                LTAPTGCLKKHGYTVEVQFDGDVHNTMHYTNWKFIYLCIDGQCTVVEGQVNCKGIYYV
                HEGHITYFVNFTEEAKKYGTGKKWEVHAGGQVIVFPESVFSSDEISFAGIVTKLPTAN
                NTTTSNSKTCALGTSEGVRRATTSTKRPRTEPEHRNTHHPNKLLRGDSVDSVNCGVIS
                AAACTNQTRAVSCPATTPIIHLKGDANILKCLRYRLSKYKQLYEQVSSTWHWTCTDGK
                HKNAIVTLTYISTSQRDDFLNTVKIPNTVSVSTGYMTI"
gene            3270. .3578
                /gene="E4"
CDS             <3270. .3578
                /gene="E4"
                /note="ORF E4 from bp 3270 to 3578"
                /codon_start=1
                /protein_id="AAA46949.1"
                /db_xref="GI:459915"
                /translation="LFFLNLYLAVTKYPLLGLLQSYQQPTTPPHRIPKPAPWAPVKVC
                GGRRRLLSDQEQSQSTETPTTPTSCCEATPWTVSTVGLSVQLHAQTKQGLSVVLQLHL
                "
gene            3816. .4070
                /gene="E5"
CDS             3816. .4070
                /gene="E5"
                /note="ORF E5 from bp 3804 to 4070"
                /codon_start=1
                /protein_id="AAA46954.1"
                /db_xref="GI:459920"
                /translation="MIELNISTVSIVLCFLLCFCVLLFVCLVIRPLVLSVSVYATLLL
                LIVILWVIATSPLRCFCIYVVFIYIPLFVIHTHASFLSQQ"
repeat_region   4099. .4134
polyA_signal    4138. .4143
                /note="putative"
repeat_region   4143. .4158
gene            4171. .5571
                /gene="L2"
CDS             4171. .5571
                /gene="L2"
                /note="ORF L2 from bp 4060 to bp 5571"
                /codon_start=1
                /product="minor capsid protein"
                /protein_id="AAA46955.1"
                /db_xref="GI:459921"
                /translation="MRSKRSTKRTKRASATQLYQTCKAAGTCPSDVIPKIEHTTIADQ
                ILRYGSMGVFFGGLGIGSGSGTGGRTGYVPLSTRPSTVSEASIPIRPPVSIDPVGPLD
                PSIVSLVEESGIVDVGAPAPIPHPPTTSGFDIATTADTTPAILDVTSVSTHENPTFTD
                PSVLQPPTPAETSGHLLLSSSSISTHNYEEIPMDTFIVSTNNENITSSTPIPGVRRPA
                RLGLYSKATQQVKVIDPTFLSAPKQLITYENPAYETVNAEESLYFSNTSHNIAPDPDF
                LDIIALHRPALTSRRNTVRYSRLGNKQTLRTRSGATIGARVHYYYDISSINPAGESIE
                MQPLGASATTTSTLNDGLYDIYADTDFTVDTPATHNVSPSTAVQSTSAVSAYVPTNTT
                VPLSTGFDIPIFSGPDVPIEHAPTQVFPFPLAPTTPQVSIFVDGGDFYLHPSYYMLKR
                RRKRVSYFFTDVSVAA"
gene            5552. .7066
                /gene="L1"
CDS             5552. .7066
                /gene="L1"
                /note="ORF L1 from bp 5516 to  7066"
                /codon_start=1
                /product="major capsid protein"
                /protein_id="AAA46956.1"
                /db_xref="GI:459922"
                /translation="MSLWRPSEATVYLPPVPVSKVVSTDEYVTRTNIYYHAGSARLLT
                VGHPYYSIPKSDNPKKIVVPKVSGLQYRVFRVRLPDPNKFGFPDTSFYNPETQRLVWA
                CVGLEVGRGQPLGVGISGHPLLNKFDDTENSNRYAGGPGTDNRECISMDYKQTQLCLL
                GCKPPIGEHWGKGSPCSNNAITPGDCPPLELKNSVIQDGDMVDTGFGAMDFTALQDTK
                SNVPLDICNSICKYPDYLKMVAEPYGDTLFFYLRREQMFVRHFFNRSGTVGESVPTDL
                YIKGSGSTATLANSTYFPTPSGSMVTSDAQIFNKPYWMQRAQGHNNGICWGNQLFVTV
                VDTTRSTNMSVCAAIANSDTTFKSSNFKEYLRHGEEFDLQFIFQLCKITLSADIMTYI
                HSMNPAILEDWNFGLTTPPSGSLEDTYRFVTSQAITCQKTAPQKPKEDPFKDYVFWEV
                NLKEKFSADLDQFPLGRKFLLQAGYRARPKFKAGKRSAPSASTTTPAKRKKTKK"
polyA_signal    7227. .7231
                /note="putative"
repeat_region   7291. .7302
repeat_region   7314. .7333
                /note="putative"
protein_bind    7406. .7420
                /standard_name="glucocorticoid responsive element"
                /bound_moiety="hormone receptor"
protein_bind    7477. .7488
                /function="gene transcription"
                /bound_moiety="E2"
enhancer        7542. .7549
                /standard_name="keratinocyte-dependant enhancer"
protein_bind    7868. .7879
                /function="gene transcription"
                /bound_moiety="E2"
BASE COUNT      2528 a   1364 c   1572 g   2448 t
ORIGIN

PPH31A  Length: 7912  November 28, 2001 14:10  Type: N  Check: 5866  ..

Initial Score    =     10  Optimized Score  =     10  Significance = -0.86
Residue Identity =    45%  Matches          =     10  Mismatches   =    12
Gaps             =      0  Conservative Substitutions               =     0

                                                      X       10        20 X
                                                      TGACTGTGAACGTTCGAGATGA
                                                        ||||    |    |  ||||
   TGGAGAAACACCTACGTTGCAAGACTATGTGTTAGATTTGCAACCTGAGGCAACTGACCTCCACTGTTATGA
      570       580       590       600       610     X 620       630      X

   GCAATTACCCGACAGCTCAGATGAGGAGGATGTCATAGACAGTCCAGCTG
    640       650       660       670       680
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: November 29, 2001, 13:50:07 ; Search time 158.03 Seconds
(without alignments)
43.401 Million cell updates/sec

Title: FRAG1
Perfect score: 8
Sequence: 1 AACGTTCG 8

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1084414

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_1101:*
1: /SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	8	100.0	10	22	AAF34842	Yeast NORF gene SA
	2	8	100.0	12	20	AAZ32399	M13mp19 genome oli
	3	8	100.0	13	8	AAN70417	Oligonucleotide fo
	4	8	100.0	14	20	AAV55798	Immunostimulatory
c	5	8	100.0	15	19	AAT94072	DNA methyltransfer
c	6	8	100.0	15	19	AAT94073	DNA methyltransfer
c	7	8	100.0	15	19	AAT94074	DNA methyltransfer
c	8	8	100.0	15	19	AAT94075	DNA methyltransfer
c	9	8	100.0	15	19	AAT94076	DNA methyltransfer
c	10	8	100.0	15	19	AAT94077	DNA methyltransfer
	11	8	100.0	16	20	AAZ32398	M13mp19 genome oli
	12	8	100.0	16	22	AAS09611	Immunoreactive CpG
	13	8	100.0	16	22	AAC80641	Immunogenic CpG ol
c	14	8	100.0	17	19	AAT94091	DNA methyltransfer
c	15	8	100.0	17	19	AAT94087	DNA methyltransfer
c	16	8	100.0	17	19	AAT94088	DNA methyltransfer
c	17	8	100.0	17	19	AAT94089	DNA methyltransfer
c	18	8	100.0	17	19	AAT94090	DNA methyltransfer
	19	8	100.0	20	14	AAQ37809	3'-5'Sequence of L
c	20	8	100.0	20	17	AAT27174	PRLTS exon 2 antis
c	21	8	100.0	20	17	AAT27931	PRLTS gene exon 2
c	22	8	100.0	20	20	AAZ41866	IL-12 secretion in
	23	8	100.0	20	20	AAX95966	PCR primer used to
	24	8	100.0	20	20	AAX95959	PCR primer used to
c	25	8	100.0	20	20	AAX93475	PCR primer used to
c	26	8	100.0	20	20	AAX93466	PCR primer used to
	27	8	100.0	20	20	AAV80114	Oligo used in expe
c	28	8	100.0	20	21	AAZ61013	Nucleotide sequenc
c	29	8	100.0	20	21	AAZ47611	Parasitic infectio
c	30	8	100.0	20	21	AAZ47889	Immunostimulatory
c	31	8	100.0	20	21	AAZ47942	Immune remodeling
c	32	8	100.0	20	22	AAF98777	CpG immunostimulat
	33	8	100.0	22	19	AAV32079	Nucleotide sequenc
	34	8	100.0	22	20	AAX36624	ISS-ODN DY1018 nuc
c	35	8	100.0	22	20	AAV80105	Oligo used in expe
	36	8	100.0	22	20	AAV80096	Immunomodulatory o
	37	8	100.0	22	20	AAV80097	Immunomodulatory o
	38	8	100.0	22	20	AAV80102	Immunomodulatory o
	39	8	100.0	22	20	AAV80103	Immunomodulatory o
	40	8	100.0	22	21	AAC64051	Immunostimulatory
	41	8	100.0	22	21	AAA96253	Sequence of a stab
	42	8	100.0	22	21	AAA90458	CpG adjuvant oligo
	43	8	100.0	22	21	AAA14467	Immunostimulatory
	44	8	100.0	22	21	AAA38065	Immunostimulatory
	45	8	100.0	22	22	AAC82107	Oligonucleotide OD

ALIGNMENTS

RESULT 1
AAF34842
ID AAF34842 standard; DNA; 10 BP.
XX
AC AAF34842;
XX
DT 23-MAR-2001 (first entry)
XX
DE Yeast NORF gene SAGE tag oligonucleotide SEQ ID NO:1581.
XX
KW Yeast; Saccharomyces cerevisiae; characterisation; cell cycle; NORF;
KW nor previously assigned open reading frame; nonannotated ORF; SAGE;
KW serial analysis of gene expression; antifungal; tag; identification;
KW linker; PCR primer; ds.
XX
OS Saccharomyces cerevisiae.
XX
PN WO200077214-A2.
XX
PD 21-DEC-2000.
XX
PF 14-JUN-2000; 2000WO-US16223.
XX
PR 16-JUN-1999; 99US-0335032.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
XX
PI Velculescu V, Vogelstein B, Kinzler K;
XX
DR WPI; 2001-061874/07.
XX
PT Yeast gene coding sequences comprising NORF genes with serial analysis
PT of gene expression (SAGE) tags, useful for studying, monitoring and
PT affecting phases of the cell cycle -

XX
PS Example; Page 56; 419pp; English.
XX
CC The present invention describes an isolated DNA molecule comprising a
CC coding sequence of a yeast gene selected from a group of 745 NORF (not
CC previously assigned open reading frame; or nonannotated ORF) genes
CC comprising a SAGE (serial analysis of gene expression) tag. Also
CC described are: (1) a method (M1) of using NORF genes to affect the cell
CC cycle comprising administering a NORF gene whose expression varies by at
CC least 10% between any two phases of the cell cycle selected from log
CC phase, S phase and G2/M; (2) a method (M2) for screening candidate
CC antifungal drugs comprising: (a) contacting a test substance with a
CC yeast cell; and (b) monitoring expression of a NORF gene whose
CC expression varies as in M1, where a test substance which modifies the
CC expression of the yeast gene is a candidate antifungal drug; (3) a method
CC (M3) for identifying human genes which are involved in cell cycle
CC progression comprising contacting human DNA with a probe which comprises
CC at least 10 contiguous nucleotides of a NORF gene whose expression varies
CC as in M1; and (4) a method (M4) for identifying a candidate drug as a
CC member of a class of drugs having a characteristic effect on gene
CC expression in a yeast cell comprising contacting a yeast cell with a
CC candidate drug and monitoring expression in the yeast cell of at least 1
CC NORF gene whose expression is affected by the class of drugs. The NORF
CC genes may be used to study, monitor and affect phases of the cell cycle,
CC the differentially expressed genes may be used as markers of phases of
CC the cell cycle. The methods may be used to identify candidate drugs which
CC affect the cell cycle and for identification of antifungal drugs.
CC AAF33268 to AAF44064 represent SAGE tags used in the exemplification of
CC the present invention. AAF33262 to AAF33267 represent linkers and PCR
CC primers used in the SAGE method, in the exemplification of the present
CC invention.
XX
SQ Sequence 10 BP; 2 A; 2 C; 3 G; 3 T; 0 other;

Query Match 100.0%; Score 8; DB 22; Length 10;
Best Local Similarity 100.0%; Pred. No. 5.4e+03;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACGTTCG 8
||||||||
Db 1 aacgttcg 8

RESULT 2
AAZ32399
ID AAZ32399 standard; DNA; 12 BP.
XX
AC AAZ32399;
XX
DT 26-JAN-2000 (first entry)
XX
DE M13mp19 genome oligonucleotide SEQ ID NO:6.
XX
KW M13mp19; MGB; minor groove binder; hybridisation; conjugate;
KW mismatch discrimination; diagnosis; detection; primer; probe;
KW forensic analysis; ss.
XX
OS Synthetic.
OS Bacteriophage M13.
XX
PN WO9951621-A2.
XX
PD 14-OCT-1999.
XX
PF 05-APR-1999; 99WO-US07487.
XX
PR 03-APR-1998; 98US-0054832.
XX
PA (EPOC-) EPOCH PHARM INC.
XX
PI Hedgpeth J, Afonina IA, Kutyavin IV, Lukhtanov EA, Belousov ES;
PI Meyer RB;
XX
DR WPI; 1999-633727/54.
XX
PT Hybridization process using oligonucleotide primer or probe that is
PT conjugated to minor groove binder, e.g. for amplification reactions or
PT assays for mutations -
XX
PS Example 1; Page 33; 95pp; English.
XX
CC A method has been developed for hybridising two nucleic acids (NA) in
CC which at least one NA comprises a minor groove binder (MGB)-
CC oligonucleotide conjugate (A). MGB is a molecule of 150-2000 D that
CC binds in a non-intercalating manner to the minor groove of a double-
CC stranded NA. Hybridisation with (A), particularly where this is a probe
CC or primer, is used: in primer extension (amplification) reactions; to
CC identify single-nucleotide (nt) mismatches; in ligase reactions; in
CC sequencing; for analysis of gene expression and detection of mutations;
CC for detecting target nucleic acids (especially for diagnosis or
CC forensic analysis, e.g. to detect human immune deficiency virus or to
CC differentiate between its subtypes, including those that are resistant
CC to antiviral agents) and for cDNA synthesis. (A) forms hybrids with
CC complementary target sequences of very high stability, so even short
CC probes, e.g. 8-mers, are highly specific and efficient. (A) also improve
CC the discriminatory capacity of short oligonucleotides, providing better
CC detection of single-base mismatches, and the speed (more rapid annealing
CC to target) and versatility of assays are increased. Short primers are
CC easier, and less expensive, to produce. The present sequence represents
CC an oligonucleotide used in an example from the present invention.
XX
SQ Sequence 12 BP; 4 A; 2 C; 2 G; 4 T; 0 other;

Query Match 100.0%; Score 8; DB 20; Length 12;
Best Local Similarity 100.0%; Pred. No. 5.3e+03;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACGTTCG 8
||||||||
Db 5 aacgttcg 12

RESULT 3
AAN70417
ID AAN70417 standard; DNA; 13 BP.
XX
AC AAN70417;
XX
DT 16-FEB-1991 (first entry)
XX
DE Oligonucleotide forming part of human epidermal growth factor gene.
XX
KW Oligonucleotide; epidermal growth factor; fusion protein;
XX
OS Homo sapiens.
XX
PN EP234888-A.
XX
PD 02-SEP-1987.
XX
PF 20-FEB-1987; 87EP-0301490.
XX
PR 24-FEB-1986; 86US-0832337.
XX
PA (CREA-) CREATIVE BIOMOLECUL.
XX
PI Cohen CM, Crea R;
XX
DR WPI; 1987-244225/35.
XX
PT Human epidermal growth factor and analogues - prepd. from a
PT recombinant fusion protein attached through a glutamyl residue to
PT a leader.
XX

```
PS   Disclosure; page 17; 33pp; English.
XX
CC   The oligonucleotide is assembled with 25 other oligonucleotides to
CC   form the human EGF gene. This gene can be combined with other genetic
CC   elements to form the fusion protein X-Glu-EGF (X is an oligopeptide
CC   leader of up to 200 amino acids, Glu is a glutamyl residue). This
CC   protein can be selectively cleaved at the Glu residue adjacent to EGF
CC   using a Glu-specific protease without altering the Glu residues in
CC   the EGF molecule. EGF and analogues inhibit the secretion of gastric
CC   acid and promote cell growth. They are useful for wound healing and
CC   the treatment of gastric ulcers. They can also be used for the prepn.
CC   of antisera for use in immunoassays.
XX
SQ   Sequence 13 BP; 2 A; 4 C; 4 G; 3 T; 0 other;

  Query Match           100.0%;  Score 8;  DB 8;  Length 13;
  Best Local Similarity  100.0%;  Pred. No. 5.3e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
            ||||||||
Db       5 aacgttcg 12


RESULT   4
AAV55798
ID   AAV55798 standard; DNA; 14 BP.
XX
AC   AAV55798;
XX
DT   29-MAR-1999  (first entry)
XX
DE   Immunostimulatory sequence oligonucleotide inhibitor beta-gal/ISS-ODN.
XX
KW   Immunostimulatory sequence oligonucleotide; ISS-ODN; inhibitor;
KW   immunostimulatory activity; gene therapy; genetic immunisation;
KW   autoimmune disease; inflammation; microbial infection; immunotherapy; ss.
XX
OS   Synthetic.
XX
PN   WO9855609-A1.
XX
PD   10-DEC-1998.
XX
PF   05-JUN-1998;   98WO-US11391.
XX
PR   06-JUN-1997;   97US-0048793.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Ray E,  Roman M;
XX
DR   WPI; 1999-080827/07.
XX
PT   New oligonucleotide that inhibits action of immunostimulatory
PT   sequence oligonucleotides - particularly those present in gene
PT   therapy vectors or microbial pathogens, used to prolong gene therapy
PT   expression and to treat e.g. infections or autoimmune disease
XX
PS   Example 3; Page 31; 50pp; English.
XX
CC   This sequence represents an example of a immunostimulatory sequence
CC   oligodeoxynucleotide (ISS-ODN) inhibitor of the invention. The ISS-ODN
CC   sequences have a hexamer region of sequence 5'-Pu-Pu-Y-Z-Py-Py or
CC   5'-Pu-Pu-Y-Z-Py-polyPy for inhibiting immunostimulation caused by
CC   ISS-ODNs that contain a hexamer region consisting of at least one CpG
CC   motif flanked by two 5'-Pu and two 3'-Py. Pu = purine; Py = pyrimidine;
CC   Y = any natural or synthetic nucleotide other than C; Z = any natural or
CC   synthetic nucleotide, but if Y is not G or Inosine (I), then Z is G or I.
CC   The inhibitors are used to inhibit immunostimulatory activity of ISS-ODNs
CC   when this is present in (i) a recombinant expression vector (being used
CC   for gene therapy or genetic immunisation) or (ii) a microbe (particularly
CC   one in a host and associated with an autoimmune disease). Particularly
CC   the inhibitors prolong gene expression from the vector and reduce
CC   inflammation caused by microbial infection. They also modulate activity
CC   of ISS-ODNs, e.g. where these are used as adjuvants to boost an immune
CC   response, e.g. in immunotherapy, in contact with vertebrate lymphocytes
CC   or monocytes by reducing the Th1-type response and stimulating the
CC   Th2-type response to an antigen (including antigen-stimulated
CC   immunoglobulin G1 production). Prolonged expression from the gene therapy
CC   vector avoids the need for repeated treatments and re-engineering of the
CC   vector to eliminate ISS-ODNs. The inhibitors provide precise control over
CC   the effect of ISS-ODN-based adjuvants and can be used even where the
CC   existence, identity and location of the ISS-ODNs are unknown. The
CC   inhibitors are effective at very low doses.
XX
SQ   Sequence 14 BP; 4 A; 4 C; 2 G; 4 T; 0 other;

  Query Match           100.0%;  Score 8;  DB 20;  Length 14;
  Best Local Similarity  100.0%;  Pred. No. 5.3e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
            ||||||||
Db       6 aacgttcg 13


RESULT   5
AAT94072/c
ID   AAT94072 standard; DNA; 15 BP.
XX
AC   AAT94072;
XX
DT   22-MAY-1998  (first entry)
XX
DE   DNA methyltransferase inhibitor (5).
XX
KW   DNA methyltransferase; inhibitor; tumourigenesis; cancer; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   stem_loop       1..15
FT                   /*tag=  a
FT   modified_base   14
FT                   /*tag=  b
FT                   /mod_base= i
XX
PN   WO9744346-A2.
XX
PD   27-NOV-1997.
XX
PF   22-MAY-1997;   97WO-IB00879.
XX
PR   22-MAY-1996;   96US-0653954.
XX
PA   (UYMC-) UNIV MCGILL.
XX
PI   Bigey P,  Szyf M;
XX
DR   WPI; 1998-018424/02.
XX
PT   Novel DNA methyl:transferase enzyme inhibitor - useful for
PT   preventing tumourigenesis and cancer in humans
XX
PS   Claim 6; Page 11; 63pp; English.
XX
CC   The present DNA methyltransferase enzyme inhibitor can be used to
CC   prevent tumourigenesis and cancer, especially by forming a stable
CC   non-covalent complex with the DNA methyltransferase in a
CC   5-adenosylmethionine-independent manner. It can also be used as an
CC   analytical and diagnostic tool, and as a potentiator of transgenic
CC   plant and animal studies and gene therapy approaches. The use of
CC   inosine, uridine or 5'-bromo- or 5'-fluorocytosine in forming the
```

```
CC   hairpin results in a powerful mechanism-based inhibitor of DNA
CC   methyltransferase.
XX
SQ   Sequence 15 BP; 2 A; 3 C; 4 G; 5 T; 1 other;


  Query Match             100.0%;  Score 8;  DB 19;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       8 AACGTTCG 1


RESULT   6
AAT94073/c
ID   AAT94073 standard; DNA; 15 BP.
XX
AC   AAT94073;
XX
DT   22-MAY-1998  (first entry)
XX
DE   DNA methyltransferase inhibitor (6).
XX
KW   DNA methyltransferase; inhibitor; tumourigenesis; cancer; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   stem_loop       1..15
FT                   /*tag=  a
FT   modified_base   14
FT                   /*tag=  b
FT                   /mod_base= i
XX
PN   WO9744346-A2.
XX
PD   27-NOV-1997.
XX
PF   22-MAY-1997;    97WO-IB00879.
XX
PR   22-MAY-1996;    96US-0653954.
XX
PA   (UYMC-) UNIV MCGILL.
XX
PI.  Bigey P,  Szyf M;
XX
DR   WPI; 1998-018424/02.
XX
PT   Novel DNA methyl:transferase enzyme inhibitor - useful for
PT   preventing tumourigenesis and cancer in humans
XX
PS   Claim 6; Page 11; 63pp; English.
XX
CC   The present DNA methyltransferase enzyme inhibitor can be used to
CC   prevent tumourigenesis and cancer, especially by forming a stable
CC   non-covalent complex with the DNA methyltransferase in a
CC   5-adenosylmethionine-independent manner. It can also be used as an
CC   analytical and diagnostic tool, and as a potentiator of transgenic
CC   plant and animal studies and gene therapy approaches. The use of
CC   inosine, uridine or 5'-bromo- or 5'-fluorocytosine in forming the
CC   hairpin results in a powerful mechanism-based inhibitor of DNA
CC   methyltransferase.
XX
SQ   Sequence 15 BP; 2 A; 3 C; 4 G; 5 T; 1 other;


  Query Match             100.0%;  Score 8;  DB 19;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       8 AACGTTCG 1


RESULT   7
AAT94074/c
ID   AAT94074 standard; DNA; 15 BP.
XX
AC   AAT94074;
XX
DT   22-MAY-1998  (first entry)
XX
DE   DNA methyltransferase inhibitor (7).
XX
KW   DNA methyltransferase; inhibitor; tumourigenesis; cancer; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   stem_loop       1..15
FT                   /*tag=  a
FT   modified_base   14
FT                   /*tag=  b
FT                   /mod_base= i
XX
PN   WO9744346-A2.
XX
PD   27-NOV-1997.
XX
PF   22-MAY-1997;    97WO-IB00879.
XX
PR   22-MAY-1996;    96US-0653954.
XX
PA   (UYMC-) UNIV MCGILL.
XX
PI   Bigey P,  Szyf M;
XX
DR   WPI; 1998-018424/02.
XX
PT   Novel DNA methyl:transferase enzyme inhibitor - useful for
PT   preventing.tumourigenesis and cancer in humans
XX
PS   Claim 6; Page 11; 63pp; English.
XX
CC   The present DNA methyltransferase enzyme inhibitor can be used to
CC   prevent tumourigenesis and cancer, especially by forming a stable
CC   non-covalent complex with the DNA methyltransferase in a
CC   5-adenosylmethionine-independent manner. It can also be used as an
CC   analytical and diagnostic tool, and as a potentiator of transgenic
CC   plant and animal studies and gene therapy approaches. The use of
CC   inosine, uridine or 5'-bromo- or 5'-fluorocytosine in forming the
CC   hairpin results in a powerful mechanism-based inhibitor of DNA
CC   methyltransferase.
XX
SQ   Sequence 15 BP; 2 A; 3 C; 4 G; 5 T; 1 other;


  Query Match             100.0%;  Score 8;  DB 19;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       8 AACGTTCG 1


RESULT   8
AAT94075/c
ID   AAT94075 standard; DNA; 15 BP.
XX
AC   AAT94075;
XX
DT   22-MAY-1998  (first entry)
```

```
XX
DE   DNA methyltransferase inhibitor (8).
XX
KW   DNA methyltransferase; inhibitor; tumourigenesis; cancer; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   stem_loop       1..15
FT                   /*tag=  a
FT   misc_RNA        14
FT                   /*tag=  b
XX
PN   WO9744346-A2.
XX
PD   27-NOV-1997.
XX
PF   22-MAY-1997;   97WO-IB00879.
XX
PR   22-MAY-1996;   96US-0653954.
XX
PA   (UYMC-) UNIV MCGILL.
XX
PI   Bigey P,  Szyf M;
XX
DR   WPI; 1998-018424/02.
XX
PT   Novel DNA methyl:transferase enzyme inhibitor - useful for
PT   preventing tumourigenesis and cancer in humans
XX
PS   Claim 6; Page 11; 63pp; English.
XX
CC   The present DNA methyltransferase enzyme inhibitor can be used to
CC   prevent tumourigenesis and cancer, especially by forming a stable
CC   non-covalent complex with the DNA methyltransferase in a
CC   5-adenosylmethionine-independent manner. It can also be used as an
CC   analytical and diagnostic tool, and as a potentiator of transgenic
CC   plant and animal studies and gene therapy approaches. The use of
CC   inosine, uridine or 5'-bromo- or 5'-fluorocytosine in forming the
CC   hairpin results in a powerful mechanism-based inhibitor of DNA
CC   methyltransferase.
XX
SQ   Sequence 15 BP; 2 A; 3 C; 4 G; 5 T; 1 U; 0 other;

  Query Match             100.0%;  Score 8;  DB 19;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       8 AACGTTCG 1


RESULT   9
AAT94076/c
ID   AAT94076 standard; DNA; 15 BP.
XX
AC   AAT94076;
XX
DT   22-MAY-1998  (first entry)
XX
DE   DNA methyltransferase inhibitor (9).
XX
KW   DNA methyltransferase; inhibitor; tumourigenesis; cancer; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   stem_loop       1..15
FT                   /*tag=  a
FT   misc_RNA        14
FT                   /*tag=  b
XX
PN   WO9744346-A2.
XX
PD   27-NOV-1997.
XX
PF   22-MAY-1997;   97WO-IB00879.
XX
PR   22-MAY-1996;   96US-0653954.
XX
PA   (UYMC-) UNIV MCGILL.
XX
PI   Bigey P,  Szyf M;
XX
DR   WPI; 1998-018424/02.
XX
PT   Novel DNA methyl:transferase enzyme inhibitor - useful for
PT   preventing tumourigenesis and cancer in humans
XX
PS   Claim 6; Page 11; 63pp; English.
XX
CC   The present DNA methyltransferase enzyme inhibitor can be used to
CC   prevent tumourigenesis and cancer, especially by forming a stable
CC   non-covalent complex with the DNA methyltransferase in a
CC   5-adenosylmethionine-independent manner. It can also be used as an
CC   analytical and diagnostic tool, and as a potentiator of transgenic
CC   plant and animal studies and gene therapy approaches. The use of
CC   inosine, uridine or 5'-bromo- or 5'-fluorocytosine in forming the
CC   hairpin results in a powerful mechanism-based inhibitor of DNA
CC   methyltransferase.
XX
SQ   Sequence 15 BP; 2 A; 3 C; 4 G; 5 T; 1 U; 0 other;

  Query Match             100.0%;  Score 8;  DB 19;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       8 AACGTTCG 1


RESULT  10
AAT94077/c
ID   AAT94077 standard; DNA; 15 BP.
XX
AC   AAT94077;
XX
DT   22-MAY-1998  (first entry)
XX
DE   DNA methyltransferase inhibitor (10).
XX
KW   DNA methyltransferase; inhibitor; tumourigenesis; cancer; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   stem_loop       1..15
FT                   /*tag=  a
FT   misc_RNA        14
FT                   /*tag=  b
XX
PN   WO9744346-A2.
XX
PD   27-NOV-1997.
XX
PF   22-MAY-1997;   97WO-IB00879.
XX
PR   22-MAY-1996;   96US-0653954.
XX
PA   (UYMC-) UNIV MCGILL.
XX
PI   Bigey P,  Szyf M;
```

```
XX
DR   WPI; 1998-018424/02.
XX
PT   Novel DNA methyl:transferase enzyme inhibitor - useful for
PT   preventing tumourigenesis and cancer in humans
XX
PS   Claim 6; Page 11; 63pp; English.
XX
CC   The present DNA methyltransferase enzyme inhibitor can be used to
CC   prevent tumourigenesis and cancer, especially by forming a stable
CC   non-covalent complex with the DNA methyltransferase in a
CC   5-adenosylmethionine-independent manner. It can also be used as an
CC   analytical and diagnostic tool, and as a potentiator of transgenic
CC   plant and animal studies and gene therapy approaches. The use of
CC   inosine, uridine or 5'-bromo- or 5'-fluorocytosine in forming the
CC   hairpin results in a powerful mechanism-based inhibitor of DNA
CC   methyltransferase.
XX
SQ   Sequence 15 BP; 2 A; 3 C; 4 G; 5 T; 1 U; 0 other;

  Query Match           100.0%;  Score 8;  DB 19;  Length 15;
  Best Local Similarity  100.0%;  Pred. No. 5.3e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
            ||||||||
Db       8 AACGTTCG 1


RESULT  11
AAZ32398
ID   AAZ32398 standard; DNA; 16 BP.
XX
AC   AAZ32398;
XX
DT   26-JAN-2000  (first entry)
XX
DE   M13mp19 genome oligonucleotide SEQ ID NO:4.
XX
KW   M13mp19; MGB; minor groove binder; hybridisation; conjugate;
KW   mismatch discrimination; diagnosis; detection; primer; probe;
KW   forensic analysis; ss.
XX
OS   Synthetic.
OS   Bacteriophage M13.
XX
PN   WO9951621-A2.
XX
PD   14-OCT-1999.
XX
PF   05-APR-1999;   99WO-US07487.
XX
PR   03-APR-1998;   98US-0054832.
XX
PA   (EPOC-) EPOCH PHARM INC.
XX
PI   Hedgpeth J,  Afonina IA,  Kutyavin IV,  Lukhtanov EA,  Belousov ES;
PI   Meyer RB;
XX
DR   WPI; 1999-633727/54.
XX
PT   Hybridization process using oligonucleotide primer or probe that is
PT   conjugated to minor groove binder, e.g. for amplification reactions or
PT   assays for mutations  -
XX
PS   Example 1; Page 33; 95pp; English.
XX
CC   A method has been developed for hybridising two nucleic acids (NA) in
CC   which at least one NA comprises a minor groove binder (MGB)-
CC   oligonucleotide conjugate (A). MGB is a molecule of 150-2000 D that
CC   binds in a non-intercalating manner to the minor groove of a double-
CC   stranded NA. Hybridisation with (A), particularly where this is a probe
CC   or primer, is used: in primer extension (amplification) reactions; to
CC   identify single-nucleotide (nt) mismatches; in ligase reactions; in
CC   sequencing; for analysis of gene expression and detection of mutations;
CC   for detecting target nucleic acids (especially for diagnosis or
CC   forensic analysis, e.g. to detect human immune deficiency virus or to
CC   differentiate between its subtypes, including those that are resistant
CC   to antiviral agents) and for cDNA synthesis. (A) forms hybrids with
CC   complementary target sequences of very high stability, so even short
CC   probes, e.g. 8-mers, are highly specific and efficient. (A) also improve
CC   the discriminatory capacity of short oligonucleotides, providing better
CC   detection of single-base mismatches, and the speed (more rapid annealing
CC   to target) and versatility of assays are increased. Short primers are
CC   easier, and less expensive, to produce. The present sequence represents
CC   an oligonucleotide used in an example from the present invention.
XX
SQ   Sequence 16 BP; 5 A; 3 C; 4 G; 4 T; 0 other;

  Query Match           100.0%;  Score 8;  DB 20;  Length 16;
  Best Local Similarity  100.0%;  Pred. No. 5.3e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
            ||||||||
Db       5 aacgttcg 12


RESULT  12
AAS09611
ID   AAS09611 standard; DNA; 16 BP.
XX
AC   AAS09611;
XX
DT   26-SEP-2001  (first entry)
XX
DE   Immunoreactive CpG sequence-containing oligonucleotide #61.
XX
KW   CpG sequence; immune response; non-B cell activation; interferon gamma;
KW   IFN-gamma; humoral; antibody production; interleukin-6 production;
KW   therapeutic; allergy; asthma; cancer; autoimmune disorder; infection;
KW   bio-warfare; vaccine; antisense therapy; eczema; allergic rhinitis;
KW   coryza; hay fever; urticaria; hives; food allergy; atopic condition;
KW   hepatitis; human immunodeficiency virus; HIV; malaria; Francisella;
KW   lupus erythematosus; rheumatoid arthritis; multiple sclerosis;
KW   schistosomiasis; tuberculosis; acquired immunodeficiency syndrome; AIDS;
KW   Leishmania; Ebola; Anthrax; Listeria; ss.
XX
OS   Synthetic.
XX
PN   WO200151500-A1.
XX
PD   19-JUL-2001.
XX
PF   12-JAN-2001; 2001WO-US01122.
XX
PR   14-JAN-2000; 2000US-0176115.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Klinman D,  Ishii K,  Verthelyi D;
XX
DR   WPI; 2001-442129/47.
XX
PT   Oligodeoxynucleotides for inducing an immune response to treat and
PT   prevent an allergic reaction, cancer, an autoimmune disorder and
PT   symptoms resulting from exposure to bio-warfare agents, comprise
PT   multiple CpG sequences  -
XX
PS   Claim 5; Page 37; 48pp; English.
XX
CC   AAS09551-AAS09662 represent oligodeoxynucleotides (ODN) of at least 10
CC   nucleotides comprising multiple CpG sequences, where one of the CpG
CC   sequences is different from another of the multiple CpG sequences.
```

```
CC   The ODN are useful for inducing an immune response, preferably a cell-
CC   mediated immune response, involving non-B cell activation, interferon
CC   gamma (IFN-gamma) production or a humoral immune response involving B
CC   cell activation, antibody and interleukin-6 production in a host, for
CC   treating, preventing or ameliorating an allergic reaction, e.g. asthma,
CC   cancer, e.g. solid tumour cancer, a disease associated with the immune
CC   system e.g. autoimmune disorder or an immune system deficiency, infection
CC   or a symptom resulting from exposure to bio-warfare agent in a human. The
CC   induction of immune response improves the efficacy of a vaccine and is
CC   used in antisense therapy. The ODN are useful for treating, preventing or
CC   ameliorating allergic reactions, including eczema, allergic rhinitis or
CC   coryza, hay fever, bronchial asthma, urticaria (hives), food allergies
CC   and other atopic conditions, for improving the efficacy of vaccines
CC   against hepatitis A, B and C, human immunodeficiency virus (HIV) and
CC   malaria, for treating immune system deficiencies, e.g. lupus
CC   erythematosus and autoimmune diseases such as rheumatoid arthritis and
CC   multiple sclerosis, infections including Francisella, schistosomiasis,
CC   tuberculosis, acquired immunodeficiency syndrome (AIDS), Leishmania and
CC   symptoms resulting from exposure of bio-warfare agent, including Ebola,
CC   Anthrax and Listeria.
XX
SQ   Sequence 16 BP; 2 A; 2 C; 10 G; 2 T; 0 other;


  Query Match             100.0%;  Score 8;  DB 22;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       6 aacgttcg 13


RESULT  13
AAC80641
ID   AAC80641 standard; DNA; 16 BP.
XX
AC   AAC80641;
XX
DT   14-FEB-2001  (first entry)
XX
DE   Immunogenic CpG oligodeoxynucleotide, SEQ ID NO:61.
XX
KW   CpG oligodeoxynucleotide; unmethylated; antigen-presenting cell;
KW   immunogenic; cytokine release; natural killer cell; NK cell activation;
KW   cell-mediated immune response; T-cell response; humoral response;
KW   B-cell response; antibody production; immune response induction;
KW   vaccine; allergy; asthma; infection; bacterial; viral; fungal; protozoal;
KW   parasitic; tuberculosis; AIDS; autoimmune disease; lupus erythematosus;
KW   rheumatoid arthritis; multiple sclerosis; solid tumour; cancer;
KW   immune deficiency; biological warfare agent; cytostatic; antiarthritic;
KW   antimicrobial; antiallergic; protozoacide; tuberculostatic;
KW   antiasthmatic; dermatological; phosphorothioate; ss.
XX
OS   Synthetic.
XX
PN   WO200061151-A2.
XX
PD   19-OCT-2000.
XX
PF   12-APR-2000; 2000WO-US09839.
XX
PR   12-APR-1999;   99US-0128898.
XX
PA   (KLIN/) KLINMAN D.
PA   (ISHI/) ISHII K.
PA   (VERT/) VERTHELYI D.
XX
PI   Klinman D,  Ishii K,  Verthelyi D;
XX
DR   WPI; 2001-006880/01.
XX
PT   Novel oligonucleotides useful for the prevention and treatment of
PT   allergies, cancer, and autoimmune disorders and for ameliorating
PT   symptoms resulting from exposure to a bio-warfare agent -
XX
PS   Claim 4; Page 33; 46pp; English.
XX
CC   The invention relates to novel immunogenic CpG oligodeoxynucleotides
CC   (AAC80581-C80723). The oligonucleotide are at least 10 bases long
CC   and comprise one of the generic sequences 5'-NNNT-CpG-WNNN-3' or
CC   5'-RY-CpG-RY-3'. The central CpG motif is unmethylated, and the
CC   oligonucleotides optionally have phosphorothioate linkages which make
CC   them more resistant to degradation. The invention also relates to an
CC   oligonucleotide delivery complex comprising an oligonucleotide of the
CC   invention and a targetting agent, and a pharmaceutical composition
CC   comprising the oligonucleotide delivery complex. The oligonucleotides
CC   are able to induce either a cell-mediated (T-cell) response or a humoral
CC   (B-cell, antibody) response, with oligonucleotides of the sequence
CC   5'-RY-CpG-RY-3' being able to induce a cell-mediated response, and those
CC   of the sequence 5'-NNNT-CpG-WNNN-3' being able to induce a humoral
CC   response. It is thought that after administration, the oligonucleotide
CC   acts on antigen-presenting cells (e.g., macrophages and dendritic
CC   cells), which then release cytokines, leading to activation of natural
CC   killer (NK) cells. A cell-mediated or humoral response can then occur by
CC   activation of T- or B-cells. The induction of an immune response is
CC   useful for treating, preventing or ameliorating an allergic reaction
CC   (preferably asthma), or an infection, where an immunogenic CpG
CC   oligonucleotide is administered either alone or in combination with an
CC   anti-allergenic agent or anti-infectious agent. The allergic conditions
CC   which may be treated include eczema, allergic rhinitis, hayfever,
CC   urticaria, food allergies and other atopic conditions, and the
CC   infections which may be treated include viral, bacterial, fungal and
CC   protozoal infections such as tuberculosis, AIDS, leishmania and
CC   schistosomiasis. Immune response induction may also be used in the
CC   treatment of an autoimmune disorder (e.g., lupus erythematosus,
CC   rheumatoid arthritis and multiple sclerosis), a disease associated with
CC   immune system deficiency, and symptoms resulting from exposure to an
CC   agent of biological warfare. An immunogenic CpG oligonucleotide, either
CC   alone or in combination with an anti-cancer agent, is useful for treating
CC   solid tumour cancer. The induction of an immune response is used in
CC   antisense therapy and to improve the efficacy of a vaccine. The
CC   oligonucleotide is preferably administered to lymphocytes ex vivo,
CC   producing activated lymphocytes which are then administered to the host.
CC   The present sequence represents an immunogenic CpG oligodeoxynucleotide
CC   of the invention.
XX
SQ   Sequence 16 BP; 2 A; 2 C; 10 G; 2 T; 0 other;


  Query Match             100.0%;  Score 8;  DB 22;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       6 aacgttcg 13


RESULT  14
AAT94091/c
ID   AAT94091 standard; DNA; 17 BP.
XX
AC   AAT94091;
XX
DT   22-MAY-1998  (first entry)
XX
DE   DNA methyltransferase inhibitor (24).
XX
KW   DNA methyltransferase; inhibitor; tumourigenesis; cancer; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   stem_loop       1..17
FT                   /*tag=  a
```

```
FT   modified_base   12
FT                   /*tag= b
FT                   /note= "cytosine, inosine, uridine, 5-bromocytosine
FT                          or 5-fluorocytosine"
FT   modified_base   16
FT                   /*tag= c
FT                   /note= "cytosine, inosine, uridine, 5-bromocytosine
FT                          or 5-fluorocytosine"
XX
PN   WO9744346-A2.
XX
PD   27-NOV-1997.
XX
PF   22-MAY-1997;   97WO-IB00879.
XX
PR   22-MAY-1996;   96US-0653954.
XX
PA   (UYMC-) UNIV MCGILL.
XX
PI   Bigey P,  Szyf M;
XX
DR   WPI; 1998-018424/02.
XX
PT   Novel DNA methyl:transferase enzyme inhibitor - useful for
PT   preventing tumourigenesis and cancer in humans
XX
PS   Claim 6; Page 13; 63pp; English.
XX
CC   The present DNA methyltransferase enzyme inhibitor can be used to
CC   prevent tumourigenesis and cancer, especially by forming a stable
CC   non-covalent complex with the DNA methyltransferase in a
CC   5-adenosylmethionine-independent manner. It can also be used as an
CC   analytical and diagnostic tool, and as a potentiator of transgenic
CC   plant and animal studies and gene therapy approaches. The use of
CC   inosine, uridine or 5'-bromo- or 5'-fluorocytosine in forming the
CC   hairpin results in a powerful mechanism-based inhibitor of DNA
CC   methyltransferase.
XX
SQ   Sequence 17 BP; 2 A; 2 C; 4 G; 7 T; 2 other;
```

```
  Query Match             100.0%;  Score 8;  DB 19;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       8 AACGTTCG 1
```

```
RESULT  15
AAT94087/c
ID   AAT94087 standard; DNA; 17 BP.
XX
AC   AAT94087;
XX
DT   22-MAY-1998  (first entry)
XX
DE   DNA methyltransferase inhibitor (20).
XX
KW   DNA methyltransferase; inhibitor; tumourigenesis; cancer; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   stem_loop       1..17
FT                   /*tag= a
XX
PN   WO9744346-A2.
XX
PD   27-NOV-1997.
XX
PF   22-MAY-1997;   97WO-IB00879.
XX
PR   22-MAY-1996;   96US-0653954.
XX
PA   (UYMC-) UNIV MCGILL.
XX
PI   Bigey P,  Szyf M;
XX
DR   WPI; 1998-018424/02.
XX
PT   Novel DNA methyl:transferase enzyme inhibitor - useful for
PT   preventing tumourigenesis and cancer in humans
XX
PS   Claim 6; Page 13; 63pp; English.
XX
CC   The present DNA methyltransferase enzyme inhibitor can be used to
CC   prevent tumourigenesis and cancer, especially by forming a stable
CC   non-covalent complex with the DNA methyltransferase in a
CC   5-adenosylmethionine-independent manner. It can also be used as an
CC   analytical and diagnostic tool, and as a potentiator of transgenic
CC   plant and animal studies and gene therapy approaches. The use of
CC   inosine, uridine or 5'-bromo- or 5'-fluorocytosine in forming the
CC   hairpin results in a powerful mechanism-based inhibitor of DNA
CC   methyltransferase.
XX
SQ   Sequence 17 BP; 2 A; 4 C; 4 G; 7 T; 0 other;
```

```
  Query Match             100.0%;  Score 8;  DB 19;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       8 AACGTTCG 1
```

```
RESULT  16
AAT94088/c
ID   AAT94088 standard; DNA; 17 BP.
XX
AC   AAT94088;
XX
DT   22-MAY-1998  (first entry)
XX
DE   DNA methyltransferase inhibitor (21).
XX
KW   DNA methyltransferase; inhibitor; tumourigenesis; cancer; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   stem_loop       1..17
FT                   /*tag= a
FT   modified_base   16
FT                   /*tag= b
FT                   /mod_base= i
XX
PN   WO9744346-A2.
XX
PD   27-NOV-1997.
XX
PF   22-MAY-1997;   97WO-IB00879.
XX
PR   22-MAY-1996;   96US-0653954.
XX
PA   (UYMC-) UNIV MCGILL.
XX
PI   Bigey P,  Szyf M;
XX
DR   WPI; 1998-018424/02.
XX
PT   Novel DNA methyl:transferase enzyme inhibitor - useful for
PT   preventing tumourigenesis and cancer in humans
```

```
XX
PS   Claim 6; Page 13; 63pp; English.
XX
CC   The present DNA methyltransferase enzyme inhibitor can be used to
CC   prevent tumourigenesis and cancer, especially by forming a stable
CC   non-covalent complex with the DNA methyltransferase in a
CC   5-adenosylmethionine-independent manner. It can also be used as an
CC   analytical and diagnostic tool, and as a potentiator of transgenic
CC   plant and animal studies and gene therapy approaches. The use of
CC   inosine, uridine or 5'-bromo- or 5'-fluorocytosine in forming the
CC   hairpin results in a powerful mechanism-based inhibitor of DNA
CC   methyltransferase.
XX
SQ   Sequence 17 BP; 2 A; 3 C; 4 G; 7 T; 1 other;

  Query Match             100.0%;  Score 8;  DB 19;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       8 AACGTTCG 1


RESULT  17
AAT94089/c
ID   AAT94089 standard; DNA; 17 BP.
XX
AC   AAT94089;
XX
DT   22-MAY-1998  (first entry)
XX
DE   DNA methyltransferase inhibitor (22).
XX
KW   DNA methyltransferase; inhibitor; tumourigenesis; cancer; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   stem_loop       1..17
FT                   /*tag=  a
FT   misc_RNA        16
FT                   /*tag=  b
XX
PN   WO9744346-A2.
XX
PD   27-NOV-1997.
XX
PF   22-MAY-1997;    97WO-IB00879.
XX
PR   22-MAY-1996;    96US-0653954.
XX
PA   (UYMC-) UNIV MCGILL.
XX
PI   Bigey P,  Szyf M;
XX
DR   WPI; 1998-018424/02.
XX
PT   Novel DNA methyl:transferase enzyme inhibitor - useful for
PT   preventing tumourigenesis and cancer in humans
XX
PS   Claim 6; Page 13; 63pp; English.
XX
CC   The present DNA methyltransferase enzyme inhibitor can be used to
CC   prevent tumourigenesis and cancer, especially by forming a stable
CC   non-covalent complex with the DNA methyltransferase in a
CC   5-adenosylmethionine-independent manner. It can also be used as an
CC   analytical and diagnostic tool, and as a potentiator of transgenic
CC   plant and animal studies and gene therapy approaches. The use of
CC   inosine, uridine or 5'-bromo- or 5'-fluorocytosine in forming the
CC   hairpin results in a powerful mechanism-based inhibitor of DNA
CC   methyltransferase.
XX
SQ   Sequence 17 BP; 2 A; 3 C; 4 G; 7 T; 1 U; 0 other;

  Query Match             100.0%;  Score 8;  DB 19;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       8 AACGTTCG 1


RESULT  18
AAT94090/c
ID   AAT94090 standard; DNA; 17 BP.
XX
AC   AAT94090;
XX
DT   22-MAY-1998  (first entry)
XX
DE   DNA methyltransferase inhibitor (23).
XX
KW   DNA methyltransferase; inhibitor; tumourigenesis; cancer; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   stem_loop       1..17
FT                   /*tag=  a
FT   modified_base   12
FT                   /*tag=  b
FT                   /note= "5-fluorocytosine"
FT   modified_base   16
FT                   /*tag=  c
FT                   /note= "5-fluorocytosine"
XX
PN   WO9744346-A2.
XX
PD   27-NOV-1997.
XX
PF   22-MAY-1997;    97WO-IB00879.
XX
PR   22-MAY-1996;    96US-0653954.
XX
PA   (UYMC-) UNIV MCGILL.
XX
PI   Bigey P,  Szyf M;
XX
DR   WPI; 1998-018424/02.
XX
PT   Novel DNA methyl:transferase enzyme inhibitor - useful for
PT   preventing tumourigenesis and cancer in humans
XX
PS   Claim 6; Page 13; 63pp; English.
XX
CC   The present DNA methyltransferase enzyme inhibitor can be used to
CC   prevent tumourigenesis and cancer, especially by forming a stable
CC   non-covalent complex with the DNA methyltransferase in a
CC   5-adenosylmethionine-independent manner. It can also be used as an
CC   analytical and diagnostic tool, and as a potentiator of transgenic
CC   plant and animal studies and gene therapy approaches. The use of
CC   inosine, uridine or 5'-bromo- or 5'-fluorocytosine in forming the
CC   hairpin results in a powerful mechanism-based inhibitor of DNA
CC   methyltransferase.
XX
SQ   Sequence 17 BP; 2 A; 2 C; 4 G; 7 T; 2 other;

  Query Match             100.0%;  Score 8;  DB 19;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       1 AACGTTCG 8
           ||||||||
Db       8 AACGTTCG 1


RESULT  19
AAQ37809
ID   AAQ37809 standard; DNA; 20 BP.
XX
AC   AAQ37809;
XX
DT   11-JUL-1993  (first entry)
XX
DE   3'-5'Sequence of Listeria monocytogenes HlyA gene primer B which
DE   is complementary to the base sequence between and inclusive of
DE   882-901.
XX
KW   Hly gene; primer; probe; PCR; ss.
XX
OS   Synthetic.
XX
PN   WO9304199-A.
XX
PD   04-MAR-1993.
XX
PF   19-AUG-1992;   92WO-GB01526.
XX
PR   20-AUG-1991;   91GB-0017902.
PR   12-FEB-1992;   92GB-0002962.
XX
PA   (SCGE-) SCIENTIFIC GENERICS LTD.
XX
PI   Parton A;
XX
DR   WPI; 1993-094024/11.
XX
PT   Detecting or quantitating nucleic acids - by utilising
PT   immobilised probe- primer in cyclic amplification methods
XX
PS   Example; Page 25; 45pp; English.
XX
CC   A sample was tested for the presence of copies of the HlyA gene of
CC   Listeria monocytogenes using a surface immobilised probe-primer to
CC   detect copies of the gene, for example, probe-primer AAQ37805 or
CC   AAQ37806 which has a poly C tail at the 5' end to ensure that the
CC   capture probe is spaced away from the solid support. In order to
CC   attach the probe primer covalently by its 5' end it is first made
CC   double stranded except for the two 5' end deoxycytosine residues
CC   producing the double stranded probe primer AAQ37807.The immobilised DNA
CC   is denatured and two oligo primers are synthesised - primers A and B
CC   (see AAQ37808 and AAQ37809).
XX
SQ   Sequence 20 BP; 4 A; 8 C; 4 G; 4 T; 0 other;


  Query Match             100.0%;  Score 8;  DB 14;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 5.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       5 aacgttcg 12


RESULT  20
AAT27174/c
ID   AAT27174 standard; DNA; 20 BP.
XX
AC   AAT27174;
XX
DT   19-NOV-1996  (first entry)
XX
DE   PRLTS exon 2 antisense primer binds to bases 281-300.
XX
KW   Human; platelet-derived growth factor; chromosome 8; deletion; PRLTS;
KW   liver cancer; liver non-small cell cancer; hepatocellular carcinoma;
KW   PDGF receptor beta-like tumour suppressor protein; colon cancer; ss.
XX
OS   Synthetic.
XX
PN   JP08092291-A.
XX
PD   09-APR-1996.
XX
PF   06-JUN-1995;   95JP-0139111.
XX
PR   29-JUL-1994;   94JP-0178131.
XX
PA   (EISA ) EISAI CO LTD.
PA   (GANK-) ZH GAN KENKYUKAI.
XX
DR   WPI; 1996-236101/24.
XX
PT   PRLTS protein and a DNA encoding it - used in the detection and
PT   treatment of cancer
XX
PS   Example 9; Page 15; 18pp; Japanese.
XX
CC   The sequences given in AAT27173-84 are primers which amplify the genomic
CC   DNA fragments given in AAT27167-72 which contain the human platelet-
CC   derived growth factor receptor beta-like tumour suppressor protein
CC   (PRLTS) gene exons.  PRLTS is the product of a gene on chromosome 8
CC   present in a region commonly deleted in cases of liver cancer, liver
CC   non-small cell cancer and colon cancer.  The PRLTS protein can be used
CC   as a research agent, a detecting and diagnosing reagent and in the
CC   treatment of cancer.
XX
SQ   Sequence 20 BP; 6 A; 5 C; 7 G; 2 T; 0 other;


  Query Match             100.0%;  Score 8;  DB 17;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 5.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      17 AACGTTCG 10


RESULT  21
AAT27931/c
ID   AAT27931 standard; DNA; 20 BP.
XX
AC   AAT27931;
XX
DT   24-AUG-1996  (first entry)
XX
DE   PRLTS gene exon 2 PCR antisense primer.
XX
KW   Platelet-derived growth factor receptor beta-like tumour suppressor;
KW   PDGF; PRLTS; carcinogenesis; tumorigenesis; lung cancer;
KW   hepatocellular carcinoma; colorectal cancer; diagnosis; primer; PCR;
KW   polymerase chain reaction; single strand conformation polymorphism;
KW   SSCP; ss.
XX
OS   Synthetic.
XX
PN   EP714981-A2.
XX
PD   05-JUN-1996.
XX
PF   26-JUL-1995;   95EP-0111769.
XX
PR   29-JUL-1994;   94JP-0178131.
XX
PA   (CANC-) CANCER INST.
```

```
PA   (EISA ) EISAI CO LTD.
XX
PI   Fujiwara Y,  Nakamura Y;
XX
DR   WPI; 1996-269714/28.
XX
PT   PDGF-receptor beta-like tumour suppressor protein - for detecting
PT   gene mutation(s) to diagnose, monitor, etc. cancers of lung, liver
PT   or colon
XX
PS   Example 10; Page 40; 49pp; English.
XX
CC   A PCR primer (AAT27930) corresponds to nucleotides 3-24 of a genomic
CC   DNA fragment (AAT27924) contg. exon 2 of the novel human platelet-
CC   derived growth factor receptor beta-like tumour suppressor (PRLTS)
CC   protein gene.  It was used with an antisense primer (AAT27931)
CC   complementary to nucleotides 281-300 for the PCR amplification of
CC   exon 2, using DNA samples from cancer patients as templates.  PCRs
CC   were also performed on the other exons of the PRLTS gene (see also
CC   AAT27932-36 and AAT30428-32).  PCR products were subjected to SSCP
CC   analysis.  3 Mutations were detected: CAC to TAC at codon 23 in
CC   DNA from a colorectal cancer patient; GCG to GTG at codon 302 in a
CC   hepatocellular carcinoma (HCC) patient; and CTTTG to CTG at codon 175
CC   in another HCC patient.
XX
SQ   Sequence 20 BP; 6 A; 5 C; 7 G; 2 T; 0 other;

  Query Match           100.0%;  Score 8;  DB 17;  Length 20;
  Best Local Similarity  100.0%;  Pred. No. 5.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      17 AACGTTCG 10


RESULT  22
AAZ41866/c
ID   AAZ41866 standard; DNA; 20 BP.
XX
AC   AAZ41866;
XX
DT   24-JAN-2000  (first entry)
XX
DE   IL-12 secretion inducing CpG oligonucleotide 11.
XX
KW   CpG oligonucleotide; phosphorothioate; interleukin-12; IL-12; secretion;
KW   human PBMC; immune response; cancer; HIV; bacterial disease; asthma;
KW   neoplastic disorder; jaagsiekte; B cell; NK cell; ss; cytokine;
KW   antigen presenting cell; infection; allergic disease.
XX
OS   Synthetic.
XX
PN   WO9951259-A2.
XX
PD   14-OCT-1999.
XX
PF   02-APR-1999;   99WO-US07335.
XX
PR   03-APR-1998;   98US-0080729.
XX
PA   (IOWA ) UNIV IOWA RES FOUND.
XX
PI   Krieg AM,  Weiner G;
XX
DR   WPI; 1999-620169/53.
XX
PT   Novel synergistic combinations of immunostimulatory oligonucleotides
PT   and immunopotentiating cytokines are useful for stimulating the immune
PT   system
XX
PS   Example 8; Page 69; 91pp; English.
XX
CC   Sequences AAZ41856-Z41949 are phosphorothioate CpG oligonucleotides
CC   which are used in the invention to induce interleukin-12 (IL-12)
CC   secretion from human PBMC. The invention comprises stimulating an immune
CC   response in a subject comprising administering to a subject exposed to an
CC   antigen, an immunopotentiating cytokine and an immunostimulatory CpG
CC   oligonucleotide to induce a synergistic antigen specific immune
CC   response. The methods are useful for treating cancer by stimulating an
CC   antigen specific immune response against a cancer antigen. The methods
CC   can also be used to treat neoplastic disorders in humans, including but
CC   not limited to: sarcoma, carcinoma, fibroma, Lymphoma, melanoma,
CC   neuroblastoma, retinoblastoma, and glioma. The methods are also useful
CC   for treating infectious diseases, e.g. viral diseases such as HIV,
CC   bacterial diseases, and fungal diseases. The methods may also be used to
CC   treat allergic diseases, e.g. asthma. The methods and compositions may
CC   also be applied to treat cancer and tumours in non human subjects,
CC   e.g. cats and dogs. Neoplasias affecting agricultural livestock may also
CC   be treated and include leukaemia, haemangiopericytoma and bovine ocular
CC   neoplasia. Chronic, infectious, contagious diseases of sheep and goats
CC   caused by the bacterium Corynebacterium pseudotuberculosis, and
CC   contagious lung tumour of sheep caused by jaagsiekte may also be
CC   treated. CpG oligonucleotides can be useful in activating B cells, NK
CC   cells, and antigen presenting cells, such as monocytes and macrophages.
CC   CpG oligonucleotides enhance antibody dependent cellular cytotoxicity and
CC   can be used as an adjuvant in conjunction with tumour antigens to
CC   protect against a tumour challenge.
XX
SQ   Sequence 20 BP; 4 A; 7 C; 3 G; 6 T; 0 other;

  Query Match           100.0%;  Score 8;  DB 20;  Length 20;
  Best Local Similarity  100.0%;  Pred. No. 5.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      17 AACGTTCG 10


RESULT  23
AAX95966
ID   AAX95966 standard; DNA; 20 BP.
XX
AC   AAX95966;
XX
DT   13-SEP-1999  (first entry)
XX
DE   PCR primer used to amplify an ORF of Chlamydia pneumoniae.
XX
KW   Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KW   sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
KW   vaccine; neutralising epitope; PCR primer; ss.
XX
OS   Synthetic.
OS   Chlamydia pneumoniae.
XX
PN   WO9927105-A2.
XX
PD   03-JUN-1999.
XX
PF   20-NOV-1998;   98WO-IB01890.
XX
PR   04-NOV-1998;   98US-0107078.
PR   21-NOV-1997;   97FR-0014673.
XX
PA   (GEST ) GENSET.
XX
PI   Griffais R;
XX
DR   WPI; 1999-357842/30.
XX
PT   Genome sequence of Chlamydia pneumoniae
XX
```

```
PS   Page 1789; Disclosure; 1912pp; English.
XX
CC   AAX91991-X97517 represent PCR primers used to amplify open reading
CC   frames and other nucleic acid sequences from the genome of
CC   Chlamydia pneumoniae (see AAX91990). C. pneumoniae causes respiratory
CC   disease such as pneumonia and bronchitis and is thought to be a
CC   contributing factor in heart disease, sarcoidosis, sinusitis, purulent
CC   otitis media, erythema nodosum or pharyngitis. The polypeptides encoded
CC   by the open reading frames of the C. pneumoniae genome (see AAY34584-
CC   AAY35879) can be used in immunogenic compositions as vaccines. Vectors
CC   containing C. pneumoniae nucleotides sequences can also be used as
CC   immunogenic compositions, especially where the vector directs the
CC   expression of a neutralising epitope of C. pneumoniae.
XX
SQ   Sequence 20 BP; 6 A; 4 C; 6 G; 4 T; 0 other;

  Query Match             100.0%;  Score 8;  DB 20;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 5.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      11 aacgttcg 18


RESULT  24
AAX95959
ID   AAX95959 standard; DNA; 20 BP.
XX
AC   AAX95959;
XX
DT   13-SEP-1999  (first entry)
XX
DE   PCR primer used to amplify an ORF of Chlamydia pneumoniae.
XX
KW   Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KW   sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
KW   vaccine; neutralising epitope; PCR primer; ss.
XX
OS   Synthetic.
OS   Chlamydia pneumoniae.
XX
PN   WO9927105-A2.
XX
PD   03-JUN-1999.
XX
PF   20-NOV-1998;   98WO-IB01890.
XX
PR   04-NOV-1998;   98US-0107078.
PR   21-NOV-1997;   97FR-0014673.
XX
PA   (GEST ) GENSET.
XX
PI   Griffais R;
XX
DR   WPI; 1999-357842/30.
XX
PT   Genome sequence of Chlamydia pneumoniae
XX
PS   Page 1788; Disclosure; 1912pp; English.
XX
CC   AAX91991-X97517 represent PCR primers used to amplify open reading
CC   frames and other nucleic acid sequences from the genome of
CC   Chlamydia pneumoniae (see AAX91990). C. pneumoniae causes respiratory
CC   disease such as pneumonia and bronchitis and is thought to be a
CC   contributing factor in heart disease, sarcoidosis, sinusitis, purulent
CC   otitis media, erythema nodosum or pharyngitis. The polypeptides encoded
CC   by the open reading frames of the C. pneumoniae genome (see AAY34584-
CC   AAY35879) can be used in immunogenic compositions as vaccines. Vectors
CC   containing C. pneumoniae nucleotides sequences can also be used as
CC   immunogenic compositions, especially where the vector directs the
CC   expression of a neutralising epitope of C. pneumoniae.
XX
SQ   Sequence 20 BP; 6 A; 4 C; 6 G; 4 T; 0 other;

  Query Match             100.0%;  Score 8;  DB 20;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 5.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      11 aacgttcg 18


RESULT  25
AAX93475/c
ID   AAX93475 standard; DNA; 20 BP.
XX
AC   AAX93475;
XX
DT   13-SEP-1999  (first entry)
XX
DE   PCR primer used to amplify an ORF of Chlamydia pneumoniae.
XX
KW   Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KW   sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
KW   vaccine; neutralising epitope; PCR primer; ss.
XX
OS   Synthetic.
OS   Chlamydia pneumoniae.
XX
PN   WO9927105-A2.
XX
PD   03-JUN-1999.
XX
PF   20-NOV-1998;   98WO-IB01890.
XX
PR   04-NOV-1998;   98US-0107078.
PR   21-NOV-1997;   97FR-0014673.
XX
PA   (GEST ) GENSET.
XX
PI   Griffais R;
XX
DR   WPI; 1999-357842/30.
XX
PT   Genome sequence of Chlamydia pneumoniae
XX
PS   Page 1594; Disclosure; 1912pp; English.
XX
CC   AAX91991-X97517 represent PCR primers used to amplify open reading
CC   frames and other nucleic acid sequences from the genome of
CC   Chlamydia pneumoniae (see AAX91990). C. pneumoniae causes respiratory
CC   disease such as pneumonia and bronchitis and is thought to be a
CC   contributing factor in heart disease, sarcoidosis, sinusitis, purulent
CC   otitis media, erythema nodosum or pharyngitis. The polypeptides encoded
CC   by the open reading frames of the C. pneumoniae genome (see AAY34584-
CC   AAY35879) can be used in immunogenic compositions as vaccines. Vectors
CC   containing C. pneumoniae nucleotides sequences can also be used as
CC   immunogenic compositions, especially where the vector directs the
CC   expression of a neutralising epitope of C. pneumoniae.
XX
SQ   Sequence 20 BP; 5 A; 7 C; 3 G; 5 T; 0 other;

  Query Match             100.0%;  Score 8;  DB 20;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 5.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       8 AACGTTCG 1
```

```
RESULT  26
AAX93466/c
ID   AAX93466 standard; DNA; 20 BP.
XX
AC   AAX93466;
XX
DT   13-SEP-1999  (first entry)
XX
DE   PCR primer used to amplify an ORF of Chlamydia pneumoniae.
XX
KW   Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KW   sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
KW   vaccine; neutralising epitope; PCR primer; ss.
XX
OS   Synthetic.
OS   Chlamydia pneumoniae.
XX
PN   WO9927105-A2.
XX
PD   03-JUN-1999.
XX
PF   20-NOV-1998;   98WO-IB01890.
XX
PR   04-NOV-1998;   98US-0107078.
PR   21-NOV-1997;   97FR-0014673.
XX
PA   (GEST ) GENSET.
XX
PI   Griffais R;
XX
DR   WPI; 1999-357842/30.
XX
PT   Genome sequence of Chlamydia pneumoniae
XX
PS   Page 1594; Disclosure; 1912pp; English.
XX
CC   AAX91991-X97517 represent PCR primers used to amplify open reading
CC   frames and other nucleic acid sequences from the genome of
CC   Chlamydia pneumoniae (see AAX91990). C. pneumoniae causes respiratory
CC   disease such as pneumonia and bronchitis and is thought to be a
CC   contributing factor in heart disease, sarcoidosis, sinusitis, purulent
CC   otitis media, erythema nodosum or pharyngitis. The polypeptides encoded
CC   by the open reading frames of the C. pneumoniae genome (see AAY34584-
CC   AAY35879) can be used in immunogenic compositions as vaccines. Vectors
CC   containing C. pneumoniae nucleotides sequences can also be used as
CC   immunogenic compositions, especially where the vector directs the
CC   expression of a neutralising epitope of C. pneumoniae.
XX
SQ   Sequence 20 BP; 5 A; 7 C; 3 G; 5 T; 0 other;

  Query Match          100.0%;  Score 8;  DB 20;  Length 20;
  Best Local Similarity  100.0%;  Pred. No. 5.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       8 AACGTTCG 1


RESULT  27
AAV80114
ID   AAV80114 standard; DNA; 20 BP.
XX
AC   AAV80114;
XX
DT   12-MAR-1999  (first entry)
XX
DE   Oligo used in experiments for stimulation of cytokine production.
XX
KW   Immunomodulatory; immunostimulatory; octanucleotide; immune regulation;
KW   ISS; cancer; allergy; asthma; hepatitis B infection; papillomavirus;
KW   human immunodeficiency virus; influenza; herpes; M. tuberculosis; ss;
KW   B. pertussis; malaria; plasmodia; Leishmania; Trypanosoma; Schistosoma.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   modified_base    8
FT                    /*tag= a
FT                    /note= "5-bromocytosine"
FT
XX
PN   WO9855495-A2.
XX
PD   10-DEC-1998.
XX
PF   05-JUN-1998;   98WO-US11578.
XX
PR   06-JUN-1997;   97US-0048793.
XX
PA   (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI   Dina D,  Roman M,  Schwartz D;
XX
DR   WPI; 1999-059898/05.
XX
PT   Immunostimulatory oligonucleotides regulate the immune system - and
PT   contain an immune-stimulating octanucleotide sequence; for treating
PT   cancer, allergic and infectious diseases
XX
PS   Example 2; Page 30; 63pp; English.
XX
CC   The invention relates to immunomodulatory oligonucleotides that comprise
CC   at least 1 immunostimulatory octanucleotide sequence (ISS) where the ISS
CC   sequences are selected from the group consisting of AACGTTCC, AACGTTCG,
CC   GACGTTCC, and GACGTTCG. The immunomodulatory sequences are used to treat
CC   patients needing immune regulation, such as those suffering from cancer,
CC   an allergic disease and asthma. They are also used to prevent infectious
CC   diseases such as influenza, herpes, hepatitis B, human immunodeficiency
CC   and papillomavirus, Hemophilus influenza, Mycobacterium tuberculosis and
CC   Bordetella pertussis, malarial plasmodia, Leishmania, Trypanosoma and
CC   Schistosoma. The immunomodulatory sequences are used to screen for human
CC   immunostimulatory activity by incubating macrophage cells and the
CC   oligonucleotide; and determining the relative amount of Th1-biased
CC   cytokines in the supernatant. Sequences AAV80104 to AAV80116 represent
CC   oligonucleotides that were tested for immunostimulatory activity. These
CC   were used in experiments for the stimulation of cytokine production and
CC   were found to lack immunostimulatory activity. The invention provides
CC   specific claimed examples (AAV80096-103) of immunomodulatory sequences.
XX
SQ   Sequence 20 BP; 4 A; 5 C; 4 G; 7 T; 0 other;

  Query Match          100.0%;  Score 8;  DB 20;  Length 20;
  Best Local Similarity  100.0%;  Pred. No. 5.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       6 aacgttcg 13


RESULT  28
AAZ61013/c
ID   AAZ61013 standard; DNA; 20 BP.
XX
AC   AAZ61013;
XX
DT   30-MAY-2000  (first entry)
XX
DE   Nucleotide sequence of an immunostimulatory CpG oligonucleotide.
XX
KW   Immunostimulatory; stereoisomer; CpG oligonucleotide; Th2; Th1; asthma;
KW   allergic reaction; allergen; cancer antigen; cancer; immunoinhibitory;
KW   inflammatory disease; inflammatory bowel disease; autoimmune disease;
KW   gingivitis; psoriasis; sepsis; ss.
```

XX
OS Synthetic.
XX
PN WO200006588-A1.
XX
PD 10-FEB-2000.
XX
PF 27-JUL-1999; 99WO-US17100.
XX
PR 27-JUL-1998; 98US-0094370.
XX
PA (IOWA) UNIV IOWA RES FOUND.
PA (CPGI-) CPG IMMUNOPHARMACEUTICALS INC.
XX
PI Krieg AM;
XX
DR WPI; 2000-195254/17.
XX
PT Immunostimulatory and immunoinhibitory stereoisomers of CpG
PT oligonucleotides useful for immunotherapy of cancer -
XX
PS Disclosure; Page 12; 88pp; English.
XX
CC AAZ60933-Z61015 represent immunostimulatory stereoisomers of CpG
CC oligonucleotides. The sequences are derived from generic nucleic
CC acid sequence, from which immunoinhibitory sequences may also be
CC derived. The immunostimulatory nucleic acids can be co-administered
CC with an antigen to induce an antigen-specific immune response. The
CC immunostimulatory nucleic acids can also be used in methods for
CC redirecting a subject's immune response from a Th2 to a Th1, for
CC treating asthma, for desensitising a subject against the occurrence
CC of an allergic reaction in response to contact with an allergen, for
CC activating an immune cell, especially a lymphocyte or a dendritic cell
CC expressing a cancer antigen or for treating cancer. The immunoinhibitory
CC nucleic acid can be used to prevent an immune response, especially where
CC the immune response in the subject is excessive due to having received
CC an immune stimulating compound. The immunoinhibitory nucleic acid can
CC be used to treat a subject having or at risk of an inflammatory disease,
CC especially inflammatory bowel disease, autoimmune disease, gingivitis,
CC psoriasis and sepsis.
XX
SQ Sequence 20 BP; 4 A; 7 C; 3 G; 6 T; 0 other;

```
  Query Match             100.0%;  Score 8;  DB 21;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 5.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
            ||||||||
Db      17 AACGTTCG 10
```

RESULT 29
AAZ47611/c
ID AAZ47611 standard; DNA; 20 BP.
XX
AC AAZ47611;
XX
DT 01-MAR-2000 (first entry)
XX
DE Parasitic infection preventing exemplary oligonucleotide SEQ ID NO:9.
XX
KW Immune system; immunostimulatory; parasitic infection; parasite;
KW CpG oligonucleotide; antigen presenting cell; natural killer cell;
KW granulocyte; malaria; helminth disease; tick; mite; ss.
XX
OS Synthetic.
XX
PN WO9956755-A1.
XX
PD 11-NOV-1999.
XX
PF 06-MAY-1999; 99WO-US09863.
XX
PR 06-MAY-1998; 98US-0084512.
XX
PA (IOWA) UNIV IOWA RES FOUND.
PA (OTTA-) OTTAWA CIVIC LOEB RES INST.
PA (USNA) US SEC OF NAVY.
XX
PI Gramzinski RA, Krieg AM, Davis HL, Hoffman SL;
XX
DR WPI; 2000-062123/05.
XX
PT Treating and preventing parasitic infections using CpG oligonucleotides
XX
PS Disclosure; Page 19; 74pp; English.
XX
CC The present invention describes a method for treating and preventing
CC parasitic infection by administration of unmethylated CpG
CC oligonucleotides. The CpG oligonucleotides are able to stimulate the
CC innate immune system via the activation of immune cells, such as antigen
CC presenting cells, natural killer cells and granulocytes. The CpG
CC oligonucleotides and the method can be used to treat and prevent
CC parasitic diseases, such as malaria, helminth diseases, tick and mites
CC in humans, animals and poultry. The oligonucleotides may be administered
CC in conjunction with parasiticides or other therapeutic compounds after
CC an organism has been diagnosed to be infected with parasites. Diseases
CC which can be treated or prevented include those caused by Plasmodium
CC falciparum, P. ovale, P. malariae, P. vivax, P. knowlesi, Babesia
CC microti, B. divergens, Trypanosoma cruzi, T. gambiense, T. rhodesiense,
CC Schistosoma mansoni, Toxoplasma gondii, Trichinella spiralis, Leishmania
CC major, L. donovani, L. braziliensis, and L. tropica. The parasite is
CC especially capable of causing malaria. The present sequence represents
CC a parasitic infection preventing exemplary oligonucleotide sequence from
CC the present invention.
XX
SQ Sequence 20 BP; 4 A; 7 C; 3 G; 6 T; 0 other;

```
  Query Match             100.0%;  Score 8;  DB 21;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 5.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
            ||||||||
Db      17 AACGTTCG 10
```

RESULT 30
AAZ47889/c
ID AAZ47889 standard; DNA; 20 BP.
XX
AC AAZ47889;
XX
DT 07-MAR-2000 (first entry)
XX
DE Immunostimulatory oligonucleotide sequence SEQ ID NO:92.
XX
KW Mucosal immunity; immunostimulatory; CpG motif; immune response;
KW antigen; allergic reaction; cancer; infectious disease; asthma; eczema;
KW allergic rhinitis; coryza; hay fever; conjunctivitis; bronchial asthma;
KW urticaria; food allergy; atopic condition; mucosal delivery; ss.
XX
OS Synthetic.
XX
PN WO9961056-A2.
XX
PD 02-DEC-1999.
XX
PF 21-MAY-1999; 99WO-US11359.
XX
PR 22-MAY-1998; 98US-0086393.
XX
PA (LOEB-) LOEB HEALTH RES INST AT OTTAWA HOSPITAL.

```
PA   (CPGI-) CPG IMMUNOPHARMACEUTICALS INC.
XX
PI   McCluskie MJ,  Davis HL;
XX
DR   WPI; 2000-062585/05.
XX
PT   Use of CpG containing oligonucleotides as adjuvants for inducing an
PT   immune response -
XX
PS   Disclosure; Page 25; 116pp; English.
XX
CC   The present invention describes a method using CpG containing
CC   oligonucleotides (ONs) as adjuvants for inducing an immune response.
CC   The method for inducing a mucosal immune response (MIR) comprises:
CC   (1) administering to a mucosal surface of a subject an ON, having a
CC   sequence including at least the formula (I); and (2) exposing the
CC   subject to an antigen to induce the MIR, where the antigen is not
CC   encoded in a nucleic acid vector: 5'X1X2CGX3X43' (I), where
CC   C and G = unmethylated, and X1, X2, X3 and X4 = nucleotides. The method
CC   can be used for treating a subject at risk of developing an allergic
CC   reaction, cancer or infectious disease. It can be used for treating
CC   asthmatic subjects, eczema, allergic rhinitis or coryza, hay fever,
CC   conjunctivitis, bronchial asthma, urticaria, food allergies or other
CC   atopic conditions. The antigen may be derived from infectious organisms
CC   such as infectious bacteria, viruses, parasites or fungi. It can be used
CC   in humans or animals, e.g. bovine, equine, feline, swine, aquatic or
CC   avian species. The ONs act as potent mucosal adjuvants to induce immune
CC   responses at both local and remote sites against an antigen
CC   administered to the mucosal tissue. Both systemic and mucosal immunity
CC   are induced by mucosal delivery of the ONs. AAZ47808 to AAZ47891
CC   represent examples of immunostimulatory oligonucleotides given in the
CC   present invention.
XX
SQ   Sequence 20 BP; 4 A; 7 C; 3 G; 6 T; 0 other;

  Query Match             100.0%;  Score 8;  DB 21;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 5.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      17 AACGTTCG 10


RESULT  31
AAZ47942/c
ID   AAZ47942 standard; DNA; 20 BP.
XX
AC   AAZ47942;
XX
DT   08-MAR-2000  (first entry)
XX
DE   Immune remodeling inducing CpG oligonucleotide SEQ ID NO:11.
XX
KW   Haematopoiesis; regulation; CpG oligonucleotide; phosphorothioate;
KW   immune remodeling; thrombopoiesis; anaemia; immune system; cancer;
KW   immune response; allergic reaction; infectious disease; asthma;
KW   thrombocytopaenia; immunohaemolytic disorder; genetic disorder;
KW   haemoglobinopathy; kidney failure; chronic inflammatory disorder;
KW   rheumatoid arthritis; ss.
XX
OS   Synthetic.
XX
PN   WO9958118-A2.
XX
PD   18-NOV-1999.
XX
PF   14-MAY-1999;   99WO-IB01285.
XX
PR   14-MAY-1998;   98US-0085516.
PR   02-FEB-1999;   99US-0241653.
XX
PA   (CPGI-) CPG IMMUNOPHARMACEUTICALS GMBH.
PA   (CPGI-) CPG IMMUNOPHARMACEUTICALS INC.
XX
PI   Wagner H,  Lipford G;
XX
DR   WPI; 2000-062261/05.
XX
PT   Use of CpG containing oligonucleotides for, e.g. inducing an
PT   antigen-specific immune response -
XX
PS   Example 1; Page 65; 116pp; English.
XX
CC   The present invention describes a method using CpG containing
CC   oligonucleotides (ONs) for regulating immune system remodeling and for
CC   regulating haematopoiesis. The method for inducing an antigen-specific
CC   immune response comprises: (1) administering an ON having a sequence
CC   including at least the formula (I); and (2) exposing the subject to an
CC   antigen at least 3 days after the ON is administered to the subject to
CC   produce an antigen-specific immune response: 5' X1CGX2 3' (I), where
CC   the ON = includes at least 8 nucleotides; C and G = unmethylated, and
CC   X1 and X2 = nucleotides. The method can be used for inducing an immune
CC   response against an antigen such as cells, cell extracts, proteins,
CC   polysaccharides, polysaccharide conjugates, lipids, glycolipids,
CC   carbohydrate, viral extracts, viruses, bacteria, fungi, parasites and
CC   allergens. It can be used in a subject at risk of developing cancer or
CC   an allergic reaction. It can also be used for treating an infectious
CC   disease, allergic diseases and asthma, as well as thrombocytopaenia
CC   which is drug-induced, due to an autoimmune disorder such as idiopathic
CC   thrombocytopenic purpura, or resulting from accidental or therapeutic
CC   radiation exposure. It can also be used for treating anaemia such as
CC   drug-induced anaemia, immunohaemolytic disorder, genetic disorders such
CC   as haemoglobinopathy and inherited haemolytic anaemia, inadequate
CC   production despite adequate iron stores, chronic disease such as kidney
CC   failure, and chronic inflammatory disorder such as rheumatoid arthritis,
CC   or anaemia resulting from accidental or therapeutic radiation exposure.
CC   AAZ47932 to AAZ48029 represent phosphorothioate CpG oligonucleotides
CC   used in the exemplification of the present invention.
XX
SQ   Sequence 20 BP; 4 A; 7 C; 3 G; 6 T; 0 other;

  Query Match             100.0%;  Score 8;  DB 21;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 5.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      17 AACGTTCG 10


RESULT  32
AAF98777/c
ID   AAF98777 standard; DNA; 20 BP.
XX
AC   AAF98777;
XX
DT   11-JUN-2001  (first entry)
XX
DE   CpG immunostimulatory nucleic acid SEQ ID NO: 48.
XX
KW   Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
KW   viral infection; phosphorothioate backbone; palindrome; cancer; ds.
XX
OS   Synthetic.
XX
PN   WO200122990-A2.
XX
PD   05-APR-2001.
XX
PF   27-SEP-2000; 2000WO-US26527.
XX
PR   27-SEP-1999;   99US-0156147.
XX
```

```
PA   (COLE-) COLEY PHARM GROUP INC.
PA   (IOWA ) UNIV IOWA RES FOUND.
XX
PI   Hartmann G,  Bratzler RL,  Krieg A;
XX
DR   WPI; 2001-290487/30.
XX
PT   Improving the efficacy of treatments involving the administration of
PT   interferon-alpha by co-administering an isolated immunostimulatory
PT   nucleic acid -
XX
PS   Disclosure; Page 21; 168pp; English.
XX
CC   The present invention describes an improvement to a method requiring the
CC   administration of interferon alpha (IFN-alpha), involving administering
CC   an immunostimulatory nucleic acid (ISNA). The sequences of a number of
CC   such nucleic acids are also provided. These may comprise oligonucleotides
CC   with phosphorothioate backbones, palindromes, or G-rich sequences. The
CC   sequences of the invention are useful in the treatment of proliferative
CC   diseases, such as cancers, and viral infections. The present sequence is
CC   an example of an immunostimulatory oligonucleotide.
XX
SQ   Sequence 20 BP; 4 A; 7 C; 3 G; 6 T; 0 other;


  Query Match             100.0%;  Score 8;  DB 22;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 5.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      17 AACGTTCG 10


RESULT  33
AAV32079
ID   AAV32079 standard; DNA; 22 BP.
XX
AC   AAV32079;
XX
DT   09-SEP-1998  (first entry)
XX
DE   Nucleotide sequence of DY1018.
XX
KW   DY1018; beta-gal; ISS-PN/IMM; antigen; immune response; antibody;
KW   immunisation; anaphylaxis; IgE; retinopathies; ss.
XX
OS   synthetic.
XX
FH   Key             Location/Qualifiers
FT   modified_base   1..22
FT                   /*tag= a
XX                   /note= "phosphothioate backbone"
PN   WO9816247-A1.
XX
PD   23-APR-1998.
XX
PF   09-OCT-1997;   97WO-US19004.
XX
PR   11-OCT-1996;   96US-0028118.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Carson DA,  Raz E,. Roman M;
XX
DR   WPI; 1998-261028/23.
XX
PT   New immunomodulatory compositions - comprising an antigen conjugated
PT   to a polynucleotide that contains an immunostimulatory sequence
XX
PS   Example 1; Page 36; 69pp; English.
XX
CC   This is the nucleotide sequence of DY1018, which is conjugated to
CC   beta-gal to form ISS-PN/IMM, comprising an immunomodulatory molecule
CC   (IMM), which comprises an antigen conjugated to a polynucleotide
CC   (PN) that contains at least one immunostimulatory nucleotide sequence
CC   (ISS).  The conjugate synergistically boost the magnitude of the host
CC   immune response against an antigen to a level greater that the host
CC   immune response to either the IMM, antigen or ISS-PN alone.  These
CC   responses to ISS-PN/IMM conjugates are particularly acute during
CC   the important early phase of the host immune response to an antigen.
CC   The ISS-PN/IMM conjugates boost both humoral (antibody) and cellular
CC   (Th1 type) immune responses of the host.  Thus, use of the method to
CC   boost the immune responsiveness of a host to subsequent challenge by a
CC   sensitising antigen without immunisation avoids the risk of
CC   Th2-mediated, immunisation-induced anaphylaxis by suppressing IgE
CC   production in response to the antigen challenge.  The conjugates can
CC   also be used to combat pathogenic infection and to stimulate
CC   therapeutic angiogenesis to treat conditions in which localised blood
CC   flow plays a significant etiological role, e.g. retinopathies.
XX
SQ   Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;


  Query Match             100.0%;  Score 8;  DB 19;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 5.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       9 aacgttcg 16


RESULT  34
AAX36624
ID   AAX36624 standard; DNA; 22 BP.
XX
AC   AAX36624;
XX
DT   09-JUL-1999  (first entry)
XX
DE   ISS-ODN DY1018 nucleotide sequence.
XX
KW   Antigen-stimulated inflammation; immunostimulatory oligonucleotide;
KW   granulocyte-mediated tissue inflammation; Th2 type immune response;
KW   immune responsiveness modulation; idiopathic hypereosiniphilic syndrome;
KW   cutaneous basophil hypersensitivity; ISS-ODN; asthma; nasal polyposis;
KW   allergic rhinitis; atopic dermatitis; allergic conjunctivitis;
XX   eosinophilic fasciitis; therapy; ss.
OS   Synthetic.
XX
PN   WO9911275-A2.
XX
PD   11-MAR-1999.
XX
PF   04-SEP-1998;   98WO-US18382.
XX
PR   05-SEP-1997;   97US-0927120.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Ray E;
XX
DR   WPI; 1999-312404/26.
XX
PT   Reducing antigen-stimulated granulocyte-mediated inflammation
XX
PS   Example 2; Page 30; 69pp; English.
XX
CC   This is the ISS-ODN DY1018 nucleotide sequence.
CC   The invention relates to a method for preventing or reducing
CC   antigen-stimulated, granulocyte-mediated tissue inflammation in a mammal,
CC   by administering an immunostimulatory oligonucleotide (ISS-ODN), where:
CC   (a) reduction in, or the absence of, a Th2 type immune response is
```

CC measured; or (b) there is a reduction or absence of other clinical signs
CC of inflammation in the host after antigen challenge. The method is used
CC to reduce or suppress granulocyte-mediated inflammation in a host tissue,
CC and to modulate the host's immune responsiveness to an antigen,
CC particularly where the subject suffers from asthma, nasal polyposis,
CC allergic rhinitis, atopic dermatitis, allergic conjunctivitis,
CC eosinophilic fasciitis, idiopathic hypereosiniphilic syndrome, or
CC cutaneous basophil hypersensitivity. Unlike prior art treatment by
CC antigen immunisation, the method is an antigen-independent method,
CC and avoids host production of both interleukin-4 (IL-4), which carries
CC risk of anaphylaxis, and IL-5 which actually encourages granulocyte
CC adhesion to endothelia.
XX
SQ Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

Query Match 100.0%; Score 8; DB 20; Length 22;
Best Local Similarity 100.0%; Pred. No. 5.2e+03;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACGTTCG 8
||||||||
Db 9 aacgttcg 16

RESULT 35
AAV80105/c
ID AAV80105 standard; DNA; 22 BP.
XX
AC AAV80105;
XX
DT 12-MAR-1999 (first entry)
XX
DE Oligo used in experiments for stimulation of cytokine production.
XX
KW Immunomodulatory; immunostimulatory; octanucleotide; immune regulation;
KW ISS: cancer; allergy; asthma; hepatitis B infection; papillomavirus;
KW human immunodeficiency virus; influenza; herpes; M. tuberculosis; ss;
KW B. pertussis; malaria; plasmodia; Leishmania; Trypanosoma; Schistosoma.
XX
OS Synthetic.
XX
PN WO9855495-A2.
XX
PD 10-DEC-1998.
XX
PF 05-JUN-1998; 98WO-US11578.
XX
PR 06-JUN-1997; 97US-0048793.
XX
PA (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI Dina D, Roman M, Schwartz D;
XX
DR WPI; 1999-059898/05.
XX
PT Immunostimulatory oligonucleotides regulate the immune system - and
PT contain an immune-stimulating octanucleotide sequence; for treating
PT cancer, allergic and infectious diseases
XX
PS Example 1; Page 29; 63pp; English.
XX
CC The invention relates to immunomodulatory oligonucleotides that comprise
CC at least 1 immunostimulatory octanucleotide sequence (ISS) where the ISS
CC sequences are selected from the group consisting of AACGTTCC, AACGTTCG,
CC GACGTTCC, and GACGTTCG. The immunomodulatory sequences are used to treat
CC patients needing immune regulation, such as those suffering from cancer,
CC an allergic disease and asthma. They are also used to prevent infectious
CC diseases such as influenza, herpes, hepatitis B, human immunodeficiency
CC and papillomavirus, Hemophilus influenza, Mycobacterium tuberculosis and
CC Bordetella pertussis, malarial plasmodia, Leishmania, Trypanosoma and
CC Schistosoma. The immunomodulatory sequences are used to screen for human
CC immunostimulatory activity by incubating macrophage cells and the
CC oligonucleotide; and determining the relative amount of Th1-biased
CC cytokines in the supernatant. Sequences AAV80104 to AAV80116 represent
CC oligonucleotides that were tested for immunostimulatory activity. These
CC were used in experiments for the stimulation of cytokine production and
CC were found to lack immunostimulatory activity. The invention provides
CC specific claimed examples (AAV80096-103) of immunomodulatory sequences.
XX
SQ Sequence 22 BP; 5 A; 7 C; 4 G; 6 T; 0 other;

Query Match 100.0%; Score 8; DB 20; Length 22;
Best Local Similarity 100.0%; Pred. No. 5.2e+03;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACGTTCG 8
||||||||
Db 14 AACGTTCG 7

RESULT 36
AAV80096
ID AAV80096 standard; DNA; 22 BP.
XX
AC AAV80096;
XX
DT 12-MAR-1999 (first entry)
XX
DE Immunomodulatory oligo comprising an ISS sequence.
XX
KW Immunomodulatory; immunostimulatory; octanucleotide; immune regulation;
KW ISS: cancer; allergy; asthma; hepatitis B infection; papillomavirus;
KW human immunodeficiency virus; influenza; herpes; M. tuberculosis; ss;
KW B. pertussis; malaria; plasmodia; Leishmania; Trypanosoma; Schistosoma.
XX
OS Synthetic.
XX
PN WO9855495-A2.
XX
PD 10-DEC-1998.
XX
PF 05-JUN-1998; 98WO-US11578.
XX
PR 06-JUN-1997; 97US-0048793.
XX
PA (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI Dina D, Roman M, Schwartz D;
XX
DR WPI; 1999-059898/05.
XX
PT Immunostimulatory oligonucleotides regulate the immune system - and
PT contain an immune-stimulating octanucleotide sequence; for treating
PT cancer, allergic and infectious diseases
XX
PS Claim 7; Page 29; 63pp; English.
XX
CC The invention relates to immunomodulatory oligonucleotides that comprise
CC at least 1 immunostimulatory octanucleotide sequence (ISS) where the ISS
CC sequences are selected from the group consisting of AACGTTCC, AACGTTCG,
CC GACGTTCC, and GACGTTCG. The immunomodulatory sequences are used to treat
CC patients needing immune regulation, such as those suffering from cancer,
CC an allergic disease and asthma. They are also used to prevent infectious
CC diseases such as influenza, herpes, hepatitis B, human immunodeficiency
CC and papillomavirus, Hemophilus influenza, Mycobacterium tuberculosis and
CC Bordetella pertussis, malarial plasmodia, Leishmania, Trypanosoma and
CC Schistosoma. The immunomodulatory sequences are used to screen for human
CC immunostimulatory activity by incubating macrophage cells and the
CC oligonucleotide; and determining the relative amount of Th1-biased
CC cytokines in the supernatant. Sequences AAV80096 to AAV80103 represent
CC specific claimed examples of such immunomodulatory oligonucleotides.
XX
SQ Sequence 22 BP; 6 A; 4 C; 7 G; 5 T; 0 other;

```
  Query Match              100.0%;  Score 8;  DB 20;  Length 22;
  Best Local Similarity    100.0%;  Pred. No. 5.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       9 aacgttcg 16


RESULT  37
AAV80097
ID   AAV80097 standard; DNA; 22 BP.
XX
AC   AAV80097;
XX
DT   12-MAR-1999  (first entry)
XX
DE   Immunomodulatory oligo comprising an ISS sequence.
XX
KW   Immunomodulatory; immunostimulatory; octanucleotide; immune regulation;
KW   ISS; cancer; allergy; asthma; hepatitis B infection; papillomavirus;
KW   human immunodeficiency virus; influenza; herpes; M. tuberculosis; ss;
KW   B. pertussis; malaria; plasmodia; Leishmania; Trypanosoma; Schistosoma.
XX
OS   Synthetic.
XX
PN   WO9855495-A2.
XX
PD   10-DEC-1998.
XX
PF   05-JUN-1998;   98WO-US11578.
XX
PR   06-JUN-1997;   97US-0048793.
XX
PA   (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI   Dina D,  Roman M,  Schwartz D;
XX
DR   WPI; 1999-059898/05.
XX
PT   Immunostimulatory oligonucleotides regulate the immune system - and
PT   contain an immune-stimulating octanucleotide sequence; for treating
PT   cancer, allergic and infectious diseases
XX
PS   Claim 5; Page 29; 63pp; English.
XX
CC   The invention relates to immunomodulatory oligonucleotides that comprise
CC   at least 1 immunostimulatory octanucleotide sequence (ISS) where the ISS
CC   sequences are selected from the group consisting of AACGTTCC, AACGTTCG,
CC   GACGTTCC, and GACGTTCG. The immunomodulatory sequences are used to treat
CC   patients needing immune regulation, such as those suffering from cancer,
CC   an allergic disease and asthma. They are also used to prevent infectious
CC   diseases such as influenza, herpes, hepatitis B, human immunodeficiency
CC   and papillomavirus, Hemophilus influenza, Mycobacterium tuberculosis and
CC   Bordetella pertussis, malarial plasmodia, Leishmania, Trypanosoma and
CC   Schistosoma. The immunomodulatory sequences are used to screen for human
CC   immunostimulatory activity by incubating macrophage cells and the
CC   oligonucleotide; and determining the relative amount of Th1-biased
CC   cytokines in the supernatant. Sequences AAV80096 to AAV80103 represent
CC   specific claimed examples of such immunomodulatory oligonucleotides.
XX
SQ   Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

  Query Match              100.0%;  Score 8;  DB 20;  Length 22;
  Best Local Similarity    100.0%;  Pred. No. 5.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       9 aacgttcg 16


RESULT  38
AAV80102
ID   AAV80102 standard; DNA; 22 BP.
XX
AC   AAV80102;
XX
DT   12-MAR-1999  (first entry)
XX
DE   Immunomodulatory oligo comprising an ISS sequence.
XX
KW   Immunomodulatory; immunostimulatory; octanucleotide; immune regulation;
KW   ISS; cancer; allergy; asthma; hepatitis B infection; papillomavirus;
KW   human immunodeficiency virus; influenza; herpes; M. tuberculosis; ss;
KW   B. pertussis; malaria; plasmodia; Leishmania; Trypanosoma; Schistosoma.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   modified_base    11
FT                    /*tag= a
FT                    /note= "5-bromocytosine"
XX
PN   WO9855495-A2.
XX
PD   10-DEC-1998.
XX
PF   05-JUN-1998;   98WO-US11578.
XX
PR   06-JUN-1997;   97US-0048793.
XX
PA   (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI   Dina D,  Roman M,  Schwartz D;
XX
DR   WPI; 1999-059898/05.
XX
PT   Immunostimulatory oligonucleotides regulate the immune system - and
PT   contain an immune-stimulating octanucleotide sequence; for treating
PT   cancer, allergic and infectious diseases
XX
PS   Claim 23; Page 30; 63pp; English.
XX
CC   The invention relates to immunomodulatory oligonucleotides that comprise
CC   at least 1 immunostimulatory octanucleotide sequence (ISS) where the ISS
CC   sequences are selected from the group consisting of AACGTTCC, AACGTTCG,
CC   GACGTTCC, and GACGTTCG. The immunomodulatory sequences are used to treat
CC   patients needing immune regulation, such as those suffering from cancer,
CC   an allergic disease and asthma. They are also used to prevent infectious
CC   diseases such as influenza, herpes, hepatitis B, human immunodeficiency
CC   and papillomavirus, Hemophilus influenza, Mycobacterium tuberculosis and
CC   Bordetella pertussis, malarial plasmodia, Leishmania, Trypanosoma and
CC   Schistosoma. The immunomodulatory sequences are used to screen for human
CC   immunostimulatory activity by incubating macrophage cells and the
CC   oligonucleotide; and determining the relative amount of Th1-biased
CC   cytokines in the supernatant. Sequences AAV80096 to AAV80103 represent
CC   specific claimed examples of such immunomodulatory oligonucleotides.
XX
SQ   Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

  Query Match              100.0%;  Score 8;  DB 20;  Length 22;
  Best Local Similarity    100.0%;  Pred. No. 5.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       9 aacgttcg 16


RESULT  39
AAV80103
ID   AAV80103 standard; DNA; 22 BP.
```

```
XX
AC   AAV80103;
XX
DT   12-MAR-1999  (first entry)
XX
DE   Immunomodulatory oligo comprising an ISS sequence.
XX
KW   Immunomodulatory; immunostimulatory; octanucleotide; immune regulation;
KW   ISS: cancer; allergy; asthma; hepatitis B infection; papillomavirus;
KW   human immunodeficiency virus; influenza; herpes; M. tuberculosis; ss;
KW   B. pertussis; malaria; plasmodia; Leishmania; Trypanosoma; Schistosoma.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   modified_base    11
FT                    /*tag=  a
FT                    /note= "5-bromocytosine"
FT
XX
PN   WO9855495-A2.
XX
PD   10-DEC-1998.
XX
PF   05-JUN-1998;   98WO-US11578.
XX
PR   06-JUN-1997;   97US-0048793.
XX
PA   (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI   Dina D,  Roman M,  Schwartz D;
XX
DR   WPI; 1999-059898/05.
XX
PT   Immunostimulatory oligonucleotides regulate the immune system - and
PT   contain an immune-stimulating octanucleotide sequence; for treating
PT   cancer, allergic and infectious diseases
XX
PS   Claim 24; Page 30; 63pp; English.
XX
CC   The invention relates to immunomodulatory oligonucleotides that comprise
CC   at least 1 immunostimulatory octanucleotide sequence (ISS) where the ISS
CC   sequences are selected from the group consisting of AACGTTCC, AACGTTCG,
CC   GACGTTCC, and GACGTTCG. The immunomodulatory sequences are used to treat
CC   patients needing immune regulation, such as those suffering from cancer,
CC   an allergic disease and asthma. They are also used to prevent infectious
CC   diseases such as influenza, herpes, hepatitis B, human immunodeficiency
CC   and papillomavirus, Hemophilus influenza, Mycobacterium tuberculosis and
CC   Bordetella pertussis, malarial plasmodia, Leishmania, Trypanosoma and
CC   Schistosoma. The immunomodulatory sequences are used to screen for human
CC   immunostimulatory activity by incubating macrophage cells and the
CC   oligonucleotide; and determining the relative amount of Th1-biased
CC   cytokines in the supernatant. Sequences AAV80096 to AAV80103 represent
CC   specific claimed examples of such immunomodulatory oligonucleotides.
XX
SQ   Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

  Query Match           100.0%;  Score 8;  DB 20;  Length 22;
  Best Local Similarity  100.0%;  Pred. No. 5.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       9 aacgttcg 16


RESULT  40
AAC64051
ID   AAC64051 standard; DNA; 22 BP.
XX
AC   AAC64051;
XX
DT   15-FEB-2001  (first entry)
XX
DE   Immunostimulatory CpG phosphorothioate oligodeoxynucleotide.
XX
KW   CpG oligodeoxynucleotide; phosphorothioate; immunostimulatory; ISS ODN;
KW   enhanced antigen presentation; antigen-presenting cell; APC;
KW   T-cell activation; tumour cell; tumour antigen; cancer immunotherapy;
KW   vaccine; ss.
XX
OS   Synthetic.
XX
PN   WO200062787-A1.
XX
PD   26-OCT-2000.
XX
PF   11-APR-2000; 2000WO-US09664.
XX
PR   15-APR-1999;   99US-0292278.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Raz E,  Martin-Orozco E;
XX
DR   WPI; 2000-679548/66.
XX
PT   Enhancing antigen-presentation capabilities of T-cells for cancer
PT   immunotherapy, by contacting cells with an immunostimulatory
PT   oligonucleotide
XX
PS   Example I; Page 18; 42pp; English.
XX
CC   The invention relates to a method of inducing activation of T-cells
CC   to respond to an antigen, comprising contacting antigen-presenting cells
CC   (APC) with an immunostimulatory oligodeoxynucleotide (ISS-ODN). The APCs
CC   thus treated have enhanced antigen presenting capabilities compared to
CC   antigen-activated APCs. APCs with enhanced antigen-presentation
CC   capabilities then present the antigen to T-cells. The method is useful
CC   for cancer immunotherapy. The ISS-ODN is used to enhance the tumour
CC   antigen presenting capacity of tumour cells, thereby inducing T-cell
CC   activation, and is therefore useful for treating tumours. Additionally,
CC   tumour cells treated with an ISS-ODN ex vivo are useful as vaccines.
CC   ISS-ODN treated APCs are induced to take up antigen through upregulation
CC   of Fc-receptor expression, to present antigen through upregulation of
CC   major histocompatibility complex (MHC) Class I and II expression and
CC   CD1d expression, to produce co-stimulatory factors (B7 and CD40), to
CC   provide cell-to-cell adhesion through upregulation of intercellular
CC   adhesion molecule (ICAM) expression, and to increase Th1 stimulatory
CC   cytokine production, all at levels greater than that achieved through
CC   contact of APC with antigen alone. The present sequence represents
CC   a phosphorothioate CpG ISS-ODN used in the exemplifications of the
CC   invention.
XX
SQ   Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

  Query Match           100.0%;  Score 8;  DB 21;  Length 22;
  Best Local Similarity  100.0%;  Pred. No. 5.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       9 aacgttcg 16


RESULT  41
AAA96253
ID   AAA96253 standard; DNA; 22 BP.
XX
AC   AAA96253;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Sequence of a stabilised oligonucleotide with antitumour activity.
XX
```

```
KW   Antitumour; immunostimulatory oligonucleotide; tumour; anaplasia;
KW   glioblastoma; medullablastoma; neuroblastoma; melanoma; carcinoma; ss.
XX
OS   Synthetic.
XX
PN   WO200056342-A2.
XX
PD   28-SEP-2000.
XX
PF   17-MAR-2000; 2000WO-FR00676.
XX
PR   19-MAR-1999;   99FR-0003433.
XX
PA   (ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.
PA   (INRM ) INST NAT SANTE & RECH MEDICALE.
XX
PI   Carpentier A;
XX
DR   WPI; 2000-602192/57.
XX
PT   Use of stabilized oligonucleotides as antitumor agents, particularly
PT   against nervous system tumors, have optimal activity and are not toxic
PT   -
XX
PS   Example 2; Page 16; 57pp; French.
XX
CC   The present sequence represents a stabilised oligonucleotide which has
CC   antitumour activity. The oligonucleotide comprises an octomer motif
CC   of the type 5'-purine-purine-CG-pyrimidine-pyrimidine-X-X-3', where
CC   the pair X-X is AT, AA, CT or TT. The oligonucleotides are
CC   immunostimulatory, and are not toxic. They may be adapted for use in
CC   animals or humans. The stabilised oligonucleotides are used for
CC   treating tumours, of any type and any degree of anaplasia, particularly
CC   human tumours in the peripheral or central nervous systems, specifically
CC   glioblastomas, medullablastomas, neuroblastomas, melanomas or carcinomas.
XX
SQ   Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

  Query Match              100.0%;  Score 8;  DB 21;  Length 22;
  Best Local Similarity    100.0%;  Pred. No. 5.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       9 aacgttcg 16


RESULT  42
AAA90458
ID   AAA90458 standard; DNA; 22 BP.
XX
AC   AAA90458;
XX
DT   10-JAN-2001  (first entry)
XX
DE   CpG adjuvant oligonucleotide, SEQ ID NO:19.
XX
KW   CpG oligonucleotide; CpG motif; adjuvant; microdroplet emulsion;
KW   microemulsion; adsorbent microparticle; vaccine; Th1 immune response;
KW   viral infection; bacterial infection; parasitic infection; HCV; HBV;
KW   hepatitis C virus; hepatitis B virus; herpes simplex virus; HSV; HIV;
KW   human immunodeficiency virus; cytomegalovirus; CMV; influenza virus;
KW   rabies virus; cholera; diphtheria; tetanus; pertussis;
KW   Helicobacter pylori; Haemophilus influenzae; malaria; ss.
XX
OS   Synthetic.
XX
PN   WO200050006-A2.
XX
PD   31-AUG-2000.
XX
PF   09-FEB-2000; 2000WO-US03331.
XX
PR   26-FEB-1999;   99US-0121858.
PR   29-JUL-1999;   99US-0146391.
PR   28-OCT-1999;   99US-0161997.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   O'Hagan D,  Ott GS,  Donnelly J,  Kazzaz J,  Ugozzoli M,  Singh M;
PI   Barackman J;
XX
DR   WPI; 2000-587123/55.
XX
PT   Microemulsion having an adsorbent surface comprising a microdroplet
PT   emulsion consisting of a metabolizable oil and an emulsifying agent
PT   which is a detergent, useful as a vaccine to treat bacterial, viral,
PT   and parasitic infection -
XX
PS   Claim 17; Page 40; 95pp; English.
XX
CC   The invention relates to a microdroplet emulsion (microemulsion) with an
CC   adsorbent surface, and which comprises a metabolisable oil and an
CC   emulsifying agent (a detergent). It also relates to a composition
CC   comprising the microemulsion and a microparticle with an adsorbent
CC   surface, where the microparticle comprises a polymer selected from a
CC   poly(alpha-hydroxy acid), a polyhydroxy butyric acid, a
CC   polycaprolactone, a polyorthoester, a polyanhydride, and a
CC   polycyanoacrylate, and a second detergent. The surface of the
CC   microparticles efficiently adsorb biologically active macromolecules such
CC   as DNA, polypeptides, antigens, hormones, pharmaceuticals, enzymes,
CC   mediators of transcription or translation, metabolic intermediates and
CC   adjuvants. Additionally, a second biologically active molecule may be
CC   encapsulated within the microparticle. The microemulsion can be used in
CC   methods of immunising a host animal, particularly a human, against a
CC   viral, bacterial or parasitic infection, and in methods of increasing a
CC   Th1 immune response. The microemulsions (having the appropriate antigens
CC   adsorbed) may be particularly used as vaccines for hepatitis C virus
CC   (HCV), hepatitis B virus (HBV), herpes simplex virus (HSV), human
CC   immunodeficiency virus (HIV), cytomegalovirus (CMV), influenza virus, and
CC   rabies virus; the bacteria which cause cholera, diphtheria, tetanus and
CC   pertussis; Helicobacter pylori and Haemophilus influenzae; and
CC   malaria-causing parasites. Sequences AAA90447-A90467 represent Th1
CC   lymphocyte stimulating oligonucleotides containing at least one CpG motif
CC   which are claimed for use as adjuvants in the compositions of the
CC   invention.
XX
SQ   Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

  Query Match              100.0%;  Score 8;  DB 21;  Length 22;
  Best Local Similarity    100.0%;  Pred. No. 5.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       9 aacgttcg 16


RESULT  43
AAA14467
ID   AAA14467 standard; DNA; 22 BP.
XX
AC   AAA14467;
XX
DT   21-AUG-2000  (first entry)
XX
DE   Immunostimulatory oligonucleotide (ISS-ODN) DY1018.
XX
KW   Immunostimulatory oligonucleotide; adjuvant; mucosal immunity;
KW   secretory immunoglobulin A production; sIgA; Th1 phenotype; ds.
XX
OS   Synthetic.
XX
PN   WO200020039-A1.
```

XX
PD 13-APR-2000.
XX
PF 15-SEP-1999; 99WO-US21203.
XX
PR 05-OCT-1998; 98US-0167039.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Raz E, Horner AA, Carson DA;
XX
DR WPI; 2000-303647/26.
XX
PT Immunostimulatory oligonucleotide adjuvant induces mucosal immunity to
PT an antigen in a mammalian host through production of secretory
PT immunoglobulin A -
XX
PS Claim 8; Page 21; 64pp; English.
XX
CC The invention relates to a method of inducing mucosal immunity to an
CC antigen in a mammalian host, including the the production of secretory
CC immunoglobulin A (sIgA). Immune protection in the mucosa (the principal
CC site of entry of most foreign antigens) is mediated by mucosa-associated
CC lymphoid tissue, epithelial and distinct B-cell, T-cell and accessory
CC cell sub-populations. The primary immune response which characterises
CC the induction of mucosal immunity to an antigen is sIgA production by
CC activated B-cells. The method comprises introducing an immunostimulatory
CC oligonucleotide (ISS-ODN) and the antigen into host mucosa, where the
CC ISS-ODN includes a core nucleotide sequence. The core nucleotide
CC sequence is 5'-Purine-Purine-C-G-Pyrimidine-Pyrimidine-3', specific
CC examples of which are AACGTT, AGCGTC and GACGTT (SEQ ID NOs 1-3). A
CC specific example of an ISS-ODN is DY1018 (AAA14467). The ISS-ODN is used
CC as an adjuvant with an antigen for stimulating mucosal immunity. The
CC level of sIgA production induced in the host is at least 3 times the
CC magnitude of sIgA production achievable in response to introduction of
CC antigen alone into the mucosal tissue and is equivalent or greater than
CC the magnitude of sIgA production achievable in response to introduction
CC of the antigen and cholera toxin adjuvant into the mucosal tissue. The
CC host immune response is stimulated to antigen specific IgA production,
CC biased towards the Th1 phenotype while antigen-induced IgE production is
CC avoided. The adjuvant has little or no known toxicity in mammals and its
CC efficacy is comparable to that of cholera toxin which is used as a
CC mucosal adjuvant. The present sequence represents the immunostimulatory
CC oligonucleotide DY1018.
XX
SQ Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

Query Match 100.0%; Score 8; DB 21; Length 22;
Best Local Similarity 100.0%; Pred. No. 5.2e+03;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACGTTCG 8
||||||||
Db 9 aacgttcg 16

RESULT 44
AAA38065
ID AAA38065 standard; DNA; 22 BP.
XX
AC AAA38065;
XX
DT 24-AUG-2000 (first entry)
XX
DE Immunostimulatory sequence (ISS) #1.
XX
KW Immunostimulatory sequence; ISS; immunomodulator; glycoprotein 120;
KW gp120; human immunodeficiency virus; HIV; immune response; infection;
KW development; ss.
XX
OS Synthetic.
XX
PN WO200021556-A1.
XX
PD 20-APR-2000.
XX
PF 08-OCT-1999; 99WO-US23677.
XX
PR 09-OCT-1998; 98US-0103733.
PR 07-OCT-1999; 99US-0415186.
XX
PA (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI Tighe H, Raz E, Schwartz D, Takabayashi K;
XX
DR WPI; 2000-317846/27.
XX
PT Anti-HIV composition comprises immunostimulatory polynucleotides and
PT HIV glycoprotein gp120 useful for modulating, stimulating an immune
PT response against HIV in an HIV infected individual -
XX
PS Claim 3; Page 16; 65pp; English.
XX
CC The present invention relates to an immunostimulatory composition
CC comprising a human immunodeficiency virus (HIV) antigen, and an
CC immunomodulatory polynucleotide comprising an immunostimulatory sequence
CC (ISS). This sequence represents an ISS that can be used in the
CC composition. An immunostimulatory composition which comprises a gp120
CC conjugated to an immunomodulatory polynucleotide, or is proximately
CC associated to it and not conjugated, is used for modulating or
CC stimulating a specific immune response against gp120 in an individual by
CC producing anti-gp120 antibodies or gp120 specific cytotoxic T cells. It
CC is also used for suppressing or delaying development of HIV infection in
CC an individual infected with HIV or an individual at risk of infection
CC with HIV, respectively. It is also used for treating an individual
CC infected with HIV in need of immune modulation.
XX
SQ Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

Query Match 100.0%; Score 8; DB 21; Length 22;
Best Local Similarity 100.0%; Pred. No. 5.2e+03;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACGTTCG 8
||||||||
Db 9 aacgttcg 16

RESULT 45
AAC82107
ID AAC82107 standard; DNA; 22 BP.
XX
AC AAC82107;
XX
DT 07-MAR-2001 (first entry)
XX
DE Oligonucleotide ODNOct DNA SEQ ID NO 2.
XX
KW Immunogenic; human immunodeficiency virus; immunostimulatory sequence;
KW ISS; beta-chemokine; anti-HIV; AIDS; Th1 immune response; primer;
KW HIV-specific cytotoxic T lymphocyte response; phosphorothioate; ss.
XX
OS Synthetic.
XX
PN WO200067787-A2.
XX
PD 16-NOV-2000.
XX
PF 05-MAY-2000; 2000WO-US12495.
XX
PR 06-MAY-1999; 99US-0132762.
PR 25-AUG-1999; 99US-0150667.
XX
PA (IMMU-) IMMUNE RESPONSE CORP.

XX
PI Moss RB;
XX
DR WPI; 2001-031804/04.
XX
PT Human immunodeficiency virus (HIV) compositions useful for immunizing
PT and inhibiting AIDS in mammals, comprises HIV devoid of outer envelope
PT protein and an immunostimulatory nucleic acid sequence -
XX
PS Example I; Page 26; 64pp; English.
XX
CC This invention describes a novel immunogenic composition (I), comprising
CC a whole-killed human immunodeficiency virus (HIV) devoid of outer
CC envelope protein gp120, an isolated nucleic acid molecule containing an
CC immunostimulatory sequence (ISS) and an adjuvant, which enhances
CC beta-chemokine levels in a mammal. The products of the invention have
CC anti-HIV activity. (I) is useful for immunizing and for inhibiting AIDS
CC in a mammal. The mammal can be a primate such as a human, (HIV
CC seronegative or seropositive humans) or a rodent, in particular the
CC primate is a pregnant mother or an infant. (I) can induce potent Th1
CC immune responses against a broad spectrum of HIV epitopes and provides a
CC strong HIV-specific cytotoxic T lymphocyte response.
XX
SQ Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

Query Match 100.0%; Score 8; DB 22; Length 22;
Best Local Similarity 100.0%; Pred. No. 5.2e+03;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACGTTCG 8
||||||||
Db 9 aacgttcg 16

Search completed: November 29, 2001, 14:51:04
Job time: 3657 sec

Mon Dec 3 08:02:31 2001

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2001, 13:48:27 ; Search time 64.43 Seconds
(without alignments)
28.121 Million cell updates/sec

Title: FRAG1
Perfect score: 8
Sequence: 1 AACGTTCG 8

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 560984

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	8	100.0	14	4	US-09-092-314-11	Sequence 11, Appl
c 2	8	100.0	15	3	US-09-206-866-5	Sequence 5, Appli
c 3	8	100.0	15	3	US-09-206-866-6	Sequence 6, Appli
c 4	8	100.0	15	3	US-09-206-866-7	Sequence 7, Appli
c 5	8	100.0	15	3	US-09-206-866-8	Sequence 8, Appli
c 6	8	100.0	15	3	US-09-206-866-9	Sequence 9, Appli
c 7	8	100.0	15	3	US-09-206-866-10	Sequence 10, Appl
c 8	8	100.0	15	4	US-09-206-866A-5	Sequence 5, Appli
c 9	8	100.0	15	4	US-09-206-866A-6	Sequence 6, Appli
c 10	8	100.0	15	4	US-09-206-866A-7	Sequence 7, Appli
c 11	8	100.0	15	4	US-09-206-866A-8	Sequence 8, Appli
c 12	8	100.0	15	4	US-09-206-866A-9	Sequence 9, Appli
c 13	8	100.0	15	4	US-09-206-866A-10	Sequence 10, Appl
c 14	8	100.0	16	3	US-09-206-866-37	Sequence 37, Appl
c 15	8	100.0	16	3	US-09-206-866-38	Sequence 38, Appl
c 16	8	100.0	16	3	US-09-206-866-39	Sequence 39, Appl
c 17	8	100.0	16	3	US-09-206-866-40	Sequence 40, Appl
c 18	8	100.0	16	3	US-09-206-866-41	Sequence 41, Appl
c 19	8	100.0	16	4	US-09-206-866A-37	Sequence 37, Appl
c 20	8	100.0	16	4	US-09-206-866A-38	Sequence 38, Appl
c 21	8	100.0	16	4	US-09-206-866A-39	Sequence 39, Appl
c 22	8	100.0	16	4	US-09-206-866A-40	Sequence 40, Appl
c 23	8	100.0	16	4	US-09-206-866A-41	Sequence 41, Appl
c 24	8	100.0	17	3	US-09-206-866-20	Sequence 20, Appl
c 25	8	100.0	17	3	US-09-206-866-21	Sequence 21, Appl
c 26	8	100.0	17	3	US-09-206-866-22	Sequence 22, Appl
c 27	8	100.0	17	3	US-09-206-866-23	Sequence 23, Appl
c 28	8	100.0	17	3	US-09-206-866-24	Sequence 24, Appl
c 29	8	100.0	17	4	US-09-206-866A-20	Sequence 20, Appl
c 30	8	100.0	17	4	US-09-206-866A-21	Sequence 21, Appl
c 31	8	100.0	17	4	US-09-206-866A-22	Sequence 22, Appl
c 32	8	100.0	17	4	US-09-206-866A-23	Sequence 23, Appl
c 33	8	100.0	17	4	US-09-206-866A-24	Sequence 24, Appl
34	8	100.0	20	1	US-08-255-892-37	Sequence 37, Appl
c 35	8	100.0	20	2	US-08-506-864A-11	Sequence 11, Appl
c 36	8	100.0	20	2	US-08-851-968-11	Sequence 11, Appl
c 37	8	100.0	20	4	US-09-286-098-11	Sequence 11, Appl
38	8	100.0	22	2	US-08-882-704A-18	Sequence 18, Appl
c 39	8	100.0	22	2	US-08-882-704A-18	Sequence 18, Appl
c 40	8	100.0	24	1	US-08-064-271-9	Sequence 9, Appli
c 41	8	100.0	24	3	US-08-930-589A-9	Sequence 9, Appli
c 42	8	100.0	26	4	US-09-129-686-16	Sequence 16, Appl
43	8	100.0	30	1	US-08-081-070-8	Sequence 8, Appli
44	8	100.0	30	1	US-08-171-389-608	Sequence 608, App
45	8	100.0	30	5	PCT-US93-12388-608	Sequence 608, App

ALIGNMENTS

RESULT 1
US-09-092-314-11
; Sequence 11, Application US/09092314
; Patent No. 6225292
; GENERAL INFORMATION:
; APPLICANT: Raz, Eyal
; APPLICANT: Roman, Mark
; TITLE OF INVENTION: Inhibitors of DNA Immunostimulatory
; TITLE OF INVENTION: Sequence Activity
; Patent No. 6225292
; FILE REFERENCE: 6510-173US1
; CURRENT APPLICATION NUMBER: US/09/092,314
; CURRENT FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/048,794
; PRIOR FILING DATE: 1997-06-06
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 14
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Oligonucleotide
US-09-092-314-11

Query Match 100.0%; Score 8; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACGTTCG 8
||||||||
Db 6 aacgttcg 13

RESULT 2
US-09-206-866-5/c
; Sequence 5, Application US/09206866A
; Patent No. 6150108
; GENERAL INFORMATION:
; APPLICANT: SZYF, Moshe
; APPLICANT: BIGEY, Pascal
; TITLE OF INVENTION: SPECIFIC INHIBITORS OF DNA METHYLTRANSFERASE
; FILE REFERENCE: 106101.200
; CURRENT APPLICATION NUMBER: US/09/206,866A
; CURRENT FILING DATE: 1998-12-08
; EARLIER APPLICATION NUMBER: US 08/653,954
; EARLIER FILING DATE: 1996-05-22
; EARLIER APPLICATION NUMBER: PCT/IB97/00879
; EARLIER FILING DATE: 1997-05-22

; EARLIER APPLICATION NUMBER: US 60/069,812
; EARLIER FILING DATE: 1997-12-17
; EARLIER APPLICATION NUMBER: US 09/194,284
; EARLIER FILING DATE: 1998-11-23
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(14)
; OTHER INFORMATION: Nucleotide 14 is n wherein n = i and i = inosine.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(15)
; OTHER INFORMATION: Nucleotides 1-15 contain c, t, a & g wherein
; OTHER INFORMATION: c=cytidine; t=thymidine; a=adenosine; g=guanosine.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(15)
; OTHER INFORMATION: m is a methyl group at the 5-position of
; OTHER INFORMATION: nucleotides 1, 5 and 10 of the cytosine portion of cytidine.
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:synthetic
; OTHER INFORMATION: construct
US-09-206-866-5

Query Match 100.0%; Score 8; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACGTTCG 8
||||||||
Db 8 AACGTTCG 1

RESULT 3
US-09-206-866-6/c
; Sequence 6, Application US/09206866A
; Patent No. 6150108
; GENERAL INFORMATION:
; APPLICANT: SZYF, Moshe
; APPLICANT: BIGEY, Pascal
; TITLE OF INVENTION: SPECIFIC INHIBITORS OF DNA METHYLTRANSFERASE
; FILE REFERENCE: 106101.200
; CURRENT APPLICATION NUMBER: US/09/206,866A
; CURRENT FILING DATE: 1998-12-08
; EARLIER APPLICATION NUMBER: US 08/653,954
; EARLIER FILING DATE: 1996-05-22
; EARLIER APPLICATION NUMBER: PCT/IB97/00879
; EARLIER FILING DATE: 1997-05-22
; EARLIER APPLICATION NUMBER: US 60/069,812
; EARLIER FILING DATE: 1997-12-17
; EARLIER APPLICATION NUMBER: US 09/194,284
; EARLIER FILING DATE: 1998-11-23
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(14)
; OTHER INFORMATION: Nucleotide 14 is n wherein n = i and i = inosine.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(15)
; OTHER INFORMATION: Nucleotides 1-15 contain c, t, a & g wherein
; OTHER INFORMATION: c=cytidine; t=thymidine; a=adenosine; g=guanosine;
; OTHER INFORMATION: m is a methyl group at the 5-position of
; OTHER INFORMATION: nucleotides 1 & 5 of the cytosine portion of
; OTHER INFORMATION: cytidine.
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:synthetic
; OTHER INFORMATION: construct
US-09-206-866-6

Query Match 100.0%; Score 8; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACGTTCG 8
||||||||
Db 8 AACGTTCG 1

RESULT 4
US-09-206-866-7/c
; Sequence 7, Application US/09206866A
; Patent No. 6150108
; GENERAL INFORMATION:
; APPLICANT: SZYF, Moshe
; APPLICANT: BIGEY, Pascal
; TITLE OF INVENTION: SPECIFIC INHIBITORS OF DNA METHYLTRANSFERASE
; FILE REFERENCE: 106101.200
; CURRENT APPLICATION NUMBER: US/09/206,866A
; CURRENT FILING DATE: 1998-12-08
; EARLIER APPLICATION NUMBER: US 08/653,954
; EARLIER FILING DATE: 1996-05-22
; EARLIER APPLICATION NUMBER: PCT/IB97/00879
; EARLIER FILING DATE: 1997-05-22
; EARLIER APPLICATION NUMBER: US 60/069,812
; EARLIER FILING DATE: 1997-12-17
; EARLIER APPLICATION NUMBER: US 09/194,284
; EARLIER FILING DATE: 1998-11-23
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(14)
; OTHER INFORMATION: Nucleotide 14 is n wherein n = i and i = inosine.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(15)
; OTHER INFORMATION: Nucleotides 1-15 contain c, t, a & g wherein
; OTHER INFORMATION: c=cytidine; t=thymidine; a=adenosine; g=guanosine;
; OTHER INFORMATION: m is a methyl group at the 5-position of
; OTHER INFORMATION: nucleotide 1 of the cytosine portion of cytidine.
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:synthetic
; OTHER INFORMATION: construct
US-09-206-866-7

Query Match 100.0%; Score 8; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACGTTCG 8
||||||||
Db 8 AACGTTCG 1

RESULT 5
US-09-206-866-8/c
; Sequence 8, Application US/09206866A
; Patent No. 6150108

```
; GENERAL INFORMATION:
;  APPLICANT: SZYF, Moshe
;  APPLICANT: BIGEY, Pascal
;  TITLE OF INVENTION: SPECIFIC INHIBITORS OF DNA METHYLTRANSFERASE
;  FILE REFERENCE: 106101.200
;  CURRENT APPLICATION NUMBER: US/09/206,866A
;  CURRENT FILING DATE: 1998-12-08
;  EARLIER APPLICATION NUMBER: US 08/653,954
;  EARLIER FILING DATE: 1996-05-22
;  EARLIER APPLICATION NUMBER: PCT/IB97/00879
;  EARLIER FILING DATE: 1997-05-22
;  EARLIER APPLICATION NUMBER: US 60/069,812
;  EARLIER FILING DATE: 1997-12-17
;  EARLIER APPLICATION NUMBER: US 09/194,284
;  EARLIER FILING DATE: 1998-11-23
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
;    LENGTH: 15
;    TYPE: DNA
;    ORGANISM: Artificial Sequence
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (1)..(14)
;    OTHER INFORMATION: Nucleotide 14 is n wherein n = u and u = uridine.
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (1)..(15)
;    OTHER INFORMATION: Nucleotides 1-15 contain c, t, a & g wherein
;    OTHER INFORMATION: c=cytidine; t=thymidine; a=adenosine; g=guanosine.
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (1)..(15)
;    OTHER INFORMATION: m is a methyl group at the 5-position of
;    OTHER INFORMATION: nucleotides 1, 5 and 10 of the cytosine portion of cytidine.
;    FEATURE:
;    OTHER INFORMATION: Description of Artificial Sequence:synthetic
;    OTHER INFORMATION: construct
US-09-206-866-8


  Query Match             100.0%;  Score 8;  DB 3;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
            ||||||||
Db       8 AACGTTCG 1


RESULT   6
US-09-206-866-9/c
; Sequence 9, Application US/09206866A
; Patent No. 6150108
; GENERAL INFORMATION:
;  APPLICANT: SZYF, Moshe
;  APPLICANT: BIGEY, Pascal
;  TITLE OF INVENTION: SPECIFIC INHIBITORS OF DNA METHYLTRANSFERASE
;  FILE REFERENCE: 106101.200
;  CURRENT APPLICATION NUMBER: US/09/206,866A
;  CURRENT FILING DATE: 1998-12-08
;  EARLIER APPLICATION NUMBER: US 08/653,954
;  EARLIER FILING DATE: 1996-05-22
;  EARLIER APPLICATION NUMBER: PCT/IB97/00879
;  EARLIER FILING DATE: 1997-05-22
;  EARLIER APPLICATION NUMBER: US 60/069,812
;  EARLIER FILING DATE: 1997-12-17
;  EARLIER APPLICATION NUMBER: US 09/194,284
;  EARLIER FILING DATE: 1998-11-23
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
;    LENGTH: 15
;    TYPE: DNA
;    ORGANISM: Artificial Sequence
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (1)..(14)
;    OTHER INFORMATION: Nucleotide 14 is n wherein n= u and u = uridine.
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (1)..(15)
;    OTHER INFORMATION: Nucleotides 1-15 contain c, t, a & g wherein
;    OTHER INFORMATION: c=cytidine; t=thymidine; a=adenosine; g=guanosine;
;    OTHER INFORMATION: m is a methyl group at the 5-position of
;    OTHER INFORMATION: nucleotides 1 & 5 of the cytosine portion of
;    OTHER INFORMATION: cytidine.
;    FEATURE:
;    OTHER INFORMATION: Description of Artificial Sequence:synthetic
;    OTHER INFORMATION: construct
US-09-206-866-9


  Query Match             100.0%;  Score 8;  DB 3;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
            ||||||||
Db       8 AACGTTCG 1


RESULT   7
US-09-206-866-10/c
; Sequence 10, Application US/09206866A
; Patent No. 6150108
; GENERAL INFORMATION:
;  APPLICANT: SZYF, Moshe
;  APPLICANT: BIGEY, Pascal
;  TITLE OF INVENTION: SPECIFIC INHIBITORS OF DNA METHYLTRANSFERASE
;  FILE REFERENCE: 106101.200
;  CURRENT APPLICATION NUMBER: US/09/206,866A
;  CURRENT FILING DATE: 1998-12-08
;  EARLIER APPLICATION NUMBER: US 08/653,954
;  EARLIER FILING DATE: 1996-05-22
;  EARLIER APPLICATION NUMBER: PCT/IB97/00879
;  EARLIER FILING DATE: 1997-05-22
;  EARLIER APPLICATION NUMBER: US 60/069,812
;  EARLIER FILING DATE: 1997-12-17
;  EARLIER APPLICATION NUMBER: US 09/194,284
;  EARLIER FILING DATE: 1998-11-23
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
;    LENGTH: 15
;    TYPE: DNA
;    ORGANISM: Artificial Sequence
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (1)..(14)
;    OTHER INFORMATION: Nucleotide 14 is n wherein n = u and u = uridine.
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (1)..(15)
;    OTHER INFORMATION: Nucleotides 1-15 contain c, t, a & g wherein
;    OTHER INFORMATION: c=cytidine; t=thymidine; a=adenosine; g=guanosine;
;    OTHER INFORMATION: m is a methyl group at the 5-position of
;    OTHER INFORMATION: nucleotide 1 of the cytosine portion of cytidine.
;    FEATURE:
;    OTHER INFORMATION: Description of Artificial Sequence:synthetic
;    OTHER INFORMATION: construct
US-09-206-866-10


  Query Match             100.0%;  Score 8;  DB 3;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+03;
```

```
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       8 AACGTTCG 1


RESULT   8
US-09-206-866A-5/c
; Sequence 5, Application US/09206866A
; Patent No. 6268137
; GENERAL INFORMATION:
;  APPLICANT: SZYF, Moshe
;  APPLICANT:  BIGEY, Pascal
;  TITLE OF INVENTION: SPECIFIC INHIBITORS OF DNA METHYLTRANSFERASE
;  FILE REFERENCE: 106101.200
;  CURRENT APPLICATION NUMBER: US/09/206,866A
;  CURRENT FILING DATE:  1998-12-08
;  PRIOR APPLICATION NUMBER: US 08/653,954
;  PRIOR FILING DATE: 1996-05-22
;  PRIOR APPLICATION NUMBER: PCT/IB97/00879
;  PRIOR FILING DATE: 1997-05-22
;  PRIOR APPLICATION NUMBER: US 60/069,812
;  PRIOR FILING DATE: 1997-12-17
;  PRIOR APPLICATION NUMBER: US 09/194,284
;  PRIOR FILING DATE: 1998-11-23
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 15
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(14)
;   OTHER INFORMATION: Nucleotide 14 is n wherein n = i and i = inosine.
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(15)
;   OTHER INFORMATION: Nucleotides 1-15 contain c, t, a & g wherein
;   OTHER INFORMATION: c=cytidine; t=thymidine; a=adenosine; g=guanosine.
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(15)
;   OTHER INFORMATION: m is a methyl group at the 5-position of
;   OTHER INFORMATION: nucleotides 1, 5 and 10 of the cytosine portion of cytidine.
;   OTHER INFORMATION: Description of Artificial Sequence:synthetic
;   OTHER INFORMATION: construct
US-09-206-866A-5


  Query Match             100.0%;  Score 8;  DB 4;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       8 AACGTTCG 1


RESULT   9
US-09-206-866A-6/c
; Sequence 6, Application US/09206866A
; Patent No. 6268137
; GENERAL INFORMATION:
;  APPLICANT: SZYF, Moshe
;  APPLICANT:  BIGEY, Pascal
;  TITLE OF INVENTION: SPECIFIC INHIBITORS OF DNA METHYLTRANSFERASE
;  FILE REFERENCE: 106101.200
;  CURRENT APPLICATION NUMBER: US/09/206,866A
;  CURRENT FILING DATE:  1998-12-08
;  PRIOR APPLICATION NUMBER: US 08/653,954
;  PRIOR FILING DATE: 1996-05-22
;  PRIOR APPLICATION NUMBER: PCT/IB97/00879
;  PRIOR FILING DATE: 1997-05-22
;  PRIOR APPLICATION NUMBER: US 60/069,812
;  PRIOR FILING DATE: 1997-12-17
;  PRIOR APPLICATION NUMBER: US 09/194,284
;  PRIOR FILING DATE: 1998-11-23
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
;   LENGTH: 15
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(14)
;   OTHER INFORMATION: Nucleotide 14 is n wherein n = i and i = inosine.
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(15)
;   OTHER INFORMATION: Nucleotides 1-15 contain c, t, a & g wherein
;   OTHER INFORMATION: c=cytidine; t=thymidine; a=adenosine; g=guanosine;
;   OTHER INFORMATION: m is a methyl group at the 5-position of
;   OTHER INFORMATION: nucleotides 1 & 5 of the cytosine portion of
;   OTHER INFORMATION: cytidine.
;   OTHER INFORMATION: Description of Artificial Sequence:synthetic
;   OTHER INFORMATION: construct
US-09-206-866A-6


  Query Match             100.0%;  Score 8;  DB 4;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       8 AACGTTCG 1


RESULT  10
US-09-206-866A-7/c
; Sequence 7, Application US/09206866A
; Patent No. 6268137
; GENERAL INFORMATION:
;  APPLICANT: SZYF, Moshe
;  APPLICANT:  BIGEY, Pascal
;  TITLE OF INVENTION: SPECIFIC INHIBITORS OF DNA METHYLTRANSFERASE
;  FILE REFERENCE: 106101.200
;  CURRENT APPLICATION NUMBER: US/09/206,866A
;  CURRENT FILING DATE:  1998-12-08
;  PRIOR APPLICATION NUMBER: US 08/653,954
;  PRIOR FILING DATE: 1996-05-22
;  PRIOR APPLICATION NUMBER: PCT/IB97/00879
;  PRIOR FILING DATE: 1997-05-22
;  PRIOR APPLICATION NUMBER: US 60/069,812
;  PRIOR FILING DATE: 1997-12-17
;  PRIOR APPLICATION NUMBER: US 09/194,284
;  PRIOR FILING DATE: 1998-11-23
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
;   LENGTH: 15
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(14)
;   OTHER INFORMATION: Nucleotide 14 is n wherein n = i and i = inosine.
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(15)
;   OTHER INFORMATION: Nucleotides 1-15 contain c, t, a & g wherein
;   OTHER INFORMATION: c=cytidine; t=thymidine; a=adenosine; g=guanosine;
;   OTHER INFORMATION: m is a methyl group at the 5-position of
;   OTHER INFORMATION: nucleotide 1 of the cytosine portion of cytidine.
;   OTHER INFORMATION: Description of Artificial Sequence:synthetic
;   OTHER INFORMATION: construct
US-09-206-866A-7
```

```
Query Match           100.0%;  Score 8;  DB 4;  Length 15;
Best Local Similarity  100.0%;  Pred. No. 1.2e+03;
Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       8 AACGTTCG 1


RESULT  11
US-09-206-866A-8/c
; Sequence 8, Application US/09206866A
; Patent No. 6268137
; GENERAL INFORMATION:
;  APPLICANT: SZYF, Moshe
;  APPLICANT:  BIGEY, Pascal
;  TITLE OF INVENTION: SPECIFIC INHIBITORS OF DNA METHYLTRANSFERASE
;  FILE REFERENCE: 106101.200
;  CURRENT APPLICATION NUMBER: US/09/206,866A
;  CURRENT FILING DATE:  1998-12-08
;  PRIOR APPLICATION NUMBER: US 08/653,954
;  PRIOR FILING DATE: 1996-05-22
;  PRIOR APPLICATION NUMBER: PCT/IB97/00879
;  PRIOR FILING DATE: 1997-05-22
;  PRIOR APPLICATION NUMBER: US 60/069,812
;  PRIOR FILING DATE: 1997-12-17
;  PRIOR APPLICATION NUMBER: US 09/194,284
;  PRIOR FILING DATE: 1998-11-23
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
;   LENGTH: 15
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(14)
;   OTHER INFORMATION: Nucleotide 14 is n wherein n = u and u = uridine.
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(15)
;   OTHER INFORMATION: Nucleotides 1-15 contain c, t, a & g wherein
;   OTHER INFORMATION: c=cytidine; t=thymidine; a=adenosine; g=guanosine.
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(15)
;   OTHER INFORMATION: m is a methyl group at the 5-position of
;   OTHER INFORMATION: nucleotides 1, 5 and 10 of the cytosine portion of cytidine.
;   OTHER INFORMATION: Description of Artificial Sequence:synthetic
;   OTHER INFORMATION: construct
US-09-206-866A-8


  Query Match           100.0%;  Score 8;  DB 4;  Length 15;
  Best Local Similarity  100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       8 AACGTTCG 1


RESULT  12
US-09-206-866A-9/c
; Sequence 9, Application US/09206866A
; Patent No. 6268137
; GENERAL INFORMATION:
;  APPLICANT: SZYF, Moshe
;  APPLICANT:  BIGEY, Pascal
;  TITLE OF INVENTION: SPECIFIC INHIBITORS OF DNA METHYLTRANSFERASE
;  FILE REFERENCE: 106101.200
;  CURRENT APPLICATION NUMBER: US/09/206,866A
;  CURRENT FILING DATE:  1998-12-08
;  PRIOR APPLICATION NUMBER: US 08/653,954
;  PRIOR FILING DATE: 1996-05-22
;  PRIOR APPLICATION NUMBER: PCT/IB97/00879
;  PRIOR FILING DATE: 1997-05-22
;  PRIOR APPLICATION NUMBER: US 60/069,812
;  PRIOR FILING DATE: 1997-12-17
;  PRIOR APPLICATION NUMBER: US 09/194,284
;  PRIOR FILING DATE: 1998-11-23
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 15
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(14)
;   OTHER INFORMATION: Nucleotide 14 is n wherein n= u and u = uridine.
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(15)
;   OTHER INFORMATION: Nucleotides 1-15 contain c, t, a & g wherein
;   OTHER INFORMATION: c=cytidine; t=thymidine; a=adenosine; g=guanosine;
;   OTHER INFORMATION: m is a methyl group at the 5-position of
;   OTHER INFORMATION: nucleotides 1 & 5 of the cytosine portion of
;   OTHER INFORMATION: cytidine.
;   OTHER INFORMATION: Description of Artificial Sequence:synthetic
;   OTHER INFORMATION: construct
US-09-206-866A-9


  Query Match           100.0%;  Score 8;  DB 4;  Length 15;
  Best Local Similarity  100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       8 AACGTTCG 1


RESULT  13
US-09-206-866A-10/c
; Sequence 10, Application US/09206866A
; Patent No. 6268137
; GENERAL INFORMATION:
;  APPLICANT: SZYF, Moshe
;  APPLICANT:  BIGEY, Pascal
;  TITLE OF INVENTION: SPECIFIC INHIBITORS OF DNA METHYLTRANSFERASE
;  FILE REFERENCE: 106101.200
;  CURRENT APPLICATION NUMBER: US/09/206,866A
;  CURRENT FILING DATE:  1998-12-08
;  PRIOR APPLICATION NUMBER: US 08/653,954
;  PRIOR FILING DATE: 1996-05-22
;  PRIOR APPLICATION NUMBER: PCT/IB97/00879
;  PRIOR FILING DATE: 1997-05-22
;  PRIOR APPLICATION NUMBER: US 60/069,812
;  PRIOR FILING DATE: 1997-12-17
;  PRIOR APPLICATION NUMBER: US 09/194,284
;  PRIOR FILING DATE: 1998-11-23
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
;   LENGTH: 15
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(14)
;   OTHER INFORMATION: Nucleotide 14 is n wherein n = u and u = uridine.
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(15)
;   OTHER INFORMATION: Nucleotides 1-15 contain c, t, a & g wherein
;   OTHER INFORMATION: c=cytidine; t=thymidine; a=adenosine; g=guanosine;
;   OTHER INFORMATION: m is a methyl group at the 5-position of
```

```
;   OTHER INFORMATION: nucleotide 1 of the cytosine portion of cytidine.
;   OTHER INFORMATION: Description of Artificial Sequence:synthetic
;   OTHER INFORMATION: construct
US-09-206-866A-10


  Query Match             100.0%;  Score 8;  DB 4;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
            ||||||||
Db       8 AACGTTCG 1


RESULT  14
US-09-206-866-37/c
; Sequence 37, Application US/09206866A
; Patent No. 6150108
; GENERAL INFORMATION:
;  APPLICANT: SZYF, Moshe
;  APPLICANT: BIGEY, Pascal
;  TITLE OF INVENTION: SPECIFIC INHIBITORS OF DNA METHYLTRANSFERASE
;  FILE REFERENCE: 106101.200
;  CURRENT APPLICATION NUMBER: US/09/206,866A
;  CURRENT FILING DATE: 1998-12-08
;  EARLIER APPLICATION NUMBER: US 08/653,954
;  EARLIER FILING DATE: 1996-05-22
;  EARLIER APPLICATION NUMBER: PCT/IB97/00879
;  EARLIER FILING DATE: 1997-05-22
;  EARLIER APPLICATION NUMBER: US 60/069,812
;  EARLIER FILING DATE: 1997-12-17
;  EARLIER APPLICATION NUMBER: US 09/194,284
;  EARLIER FILING DATE: 1998-11-23
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 37
;   LENGTH: 16
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(16)
;   OTHER INFORMATION: Nucleotides 1-16 contain c, t, a & g wherein
;   OTHER INFORMATION: c=cytidine; t=thymidine; a=adenosine; g=guanosine;
;   OTHER INFORMATION: m is a methyl group at the 5-position of
;   OTHER INFORMATION: nucleotides 1 & 5 of the cytosine porition of
;   OTHER INFORMATION: cytidine.
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:synthetic
;   OTHER INFORMATION: construct
US-09-206-866-37


  Query Match             100.0%;  Score 8;  DB 3;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
            ||||||||
Db       8 AACGTTCG 1


RESULT  15
US-09-206-866-38/c
; Sequence 38, Application US/09206866A
; Patent No. 6150108
; GENERAL INFORMATION:
;  APPLICANT: SZYF, Moshe
;  APPLICANT: BIGEY, Pascal
;  TITLE OF INVENTION: SPECIFIC INHIBITORS OF DNA METHYLTRANSFERASE
;  FILE REFERENCE: 106101.200
;  CURRENT APPLICATION NUMBER: US/09/206,866A
;  CURRENT FILING DATE: 1998-12-08
;  EARLIER APPLICATION NUMBER: US 08/653,954
;  EARLIER FILING DATE: 1996-05-22
;  EARLIER APPLICATION NUMBER: PCT/IB97/00879
;  EARLIER FILING DATE: 1997-05-22
;  EARLIER APPLICATION NUMBER: US 60/069,812
;  EARLIER FILING DATE: 1997-12-17
;  EARLIER APPLICATION NUMBER: US 09/194,284
;  EARLIER FILING DATE: 1998-11-23
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
;   LENGTH: 16
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(16)
;   OTHER INFORMATION: Nucleotides 1-16 contain c, t, a & g wherein
;   OTHER INFORMATION: c=cytidine; t=thymidine; a=adenosine; g=guanosine;
;   OTHER INFORMATION: m is a methyl group at the 5-position of
;   OTHER INFORMATION: nucleotides 1 & 5 of the cytosine portion of
;   OTHER INFORMATION: cytidine.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(15)
;   OTHER INFORMATION: Nucleotide 15 is n wherein n = i and i = inosine.
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:synthetic
;   OTHER INFORMATION: construct
US-09-206-866-38


  Query Match             100.0%;  Score 8;  DB 3;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
            ||||||||
Db       8 AACGTTCG 1


RESULT  16
US-09-206-866-39/c
; Sequence 39, Application US/09206866A
; Patent No. 6150108
; GENERAL INFORMATION:
;  APPLICANT: SZYF, Moshe
;  APPLICANT: BIGEY, Pascal
;  TITLE OF INVENTION: SPECIFIC INHIBITORS OF DNA METHYLTRANSFERASE
;  FILE REFERENCE: 106101.200
;  CURRENT APPLICATION NUMBER: US/09/206,866A
;  CURRENT FILING DATE: 1998-12-08
;  EARLIER APPLICATION NUMBER: US 08/653,954
;  EARLIER FILING DATE: 1996-05-22
;  EARLIER APPLICATION NUMBER: PCT/IB97/00879
;  EARLIER FILING DATE: 1997-05-22
;  EARLIER APPLICATION NUMBER: US 60/069,812
;  EARLIER FILING DATE: 1997-12-17
;  EARLIER APPLICATION NUMBER: US 09/194,284
;  EARLIER FILING DATE: 1998-11-23
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
;   LENGTH: 16
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(16)
;   OTHER INFORMATION: Nucleotides 1-16 contain c, t, a & g wherein
;   OTHER INFORMATION: c=cytidine; t=thymidine; a=adenosine; g=guanosine;
;   OTHER INFORMATION: m is a methyl group at the 5-position of
```

```
;   OTHER INFORMATION: nucleotides 1 & 5 of the cytosine portion of
;   OTHER INFORMATION: cytidine.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(15)
;   OTHER INFORMATION: Nucleotide 15 is n wherein n = u and u = uridine.
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:synthetic
;   OTHER INFORMATION: construct
US-09-206-866-39


  Query Match             100.0%;  Score 8;  DB 3;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
            ||||||||
Db       8 AACGTTCG 1


RESULT  17
US-09-206-866-40/c
; Sequence 40, Application US/09206866A
; Patent No. 6150108
; GENERAL INFORMATION:
;  APPLICANT: SZYF, Moshe
;  APPLICANT: BIGEY, Pascal
;  TITLE OF INVENTION: SPECIFIC INHIBITORS OF DNA METHYLTRANSFERASE
;  FILE REFERENCE: 106101.200
;  CURRENT APPLICATION NUMBER: US/09/206,866A
;  CURRENT FILING DATE: 1998-12-08
;  EARLIER APPLICATION NUMBER: US 08/653,954
;  EARLIER FILING DATE: 1996-05-22
;  EARLIER APPLICATION NUMBER: PCT/IB97/00879
;  EARLIER FILING DATE: 1997-05-22
;  EARLIER APPLICATION NUMBER: US 60/069,812
;  EARLIER FILING DATE: 1997-12-17
;  EARLIER APPLICATION NUMBER: US 09/194,284
;  EARLIER FILING DATE: 1998-11-23
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 40
;   LENGTH: 16
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(16)
;   OTHER INFORMATION: Nucleotides 1-16 contain c, t, a & g wherein
;   OTHER INFORMATION: c=cytidine; t=thymidine; a=adenosine; g=guanosine;
;   OTHER INFORMATION: m is a methyl group at the 5-position of
;   OTHER INFORMATION: nucleotides 1 & 5 of the cytosine portion of
;   OTHER INFORMATION: cytidine.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(15)
;   OTHER INFORMATION: Nucleotide 15 is n wherein n = f and f =
;   OTHER INFORMATION: 5-fluorocytosine.
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:synthetic
;   OTHER INFORMATION: construct
US-09-206-866-40


  Query Match             100.0%;  Score 8;  DB 3;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
            ||||||||
Db       8 AACGTTCG 1


RESULT  18
US-09-206-866-41/c
; Sequence 41, Application US/09206866A
; Patent No. 6150108
; GENERAL INFORMATION:
;  APPLICANT: SZYF, Moshe
;  APPLICANT: BIGEY, Pascal
;  TITLE OF INVENTION: SPECIFIC INHIBITORS OF DNA METHYLTRANSFERASE
;  FILE REFERENCE: 106101.200
;  CURRENT APPLICATION NUMBER: US/09/206,866A
;  CURRENT FILING DATE: 1998-12-08
;  EARLIER APPLICATION NUMBER: US 08/653,954
;  EARLIER FILING DATE: 1996-05-22
;  EARLIER APPLICATION NUMBER: PCT/IB97/00879
;  EARLIER FILING DATE: 1997-05-22
;  EARLIER APPLICATION NUMBER: US 60/069,812
;  EARLIER FILING DATE: 1997-12-17
;  EARLIER APPLICATION NUMBER: US 09/194,284
;  EARLIER FILING DATE: 1998-11-23
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
;   LENGTH: 16
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(16)
;   OTHER INFORMATION: Nucleotides 1-16 contain c, t, a & g wherein
;   OTHER INFORMATION: c=cytidine; t=thymidine; a=adenosine; g=guanosine;
;   OTHER INFORMATION: m is a methyl group at the 5-position of
;   OTHER INFORMATION: nucleotides 1 & 5 of the cytosine portion of
;   OTHER INFORMATION: cytidine.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(15)
;   OTHER INFORMATION: Nucleotide 15 is n wherein n = b and b = cytosine, inosine,
;   OTHER INFORMATION: uridine, 5-bromocytidine or 5-fluorouridine.
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:synthetic
;   OTHER INFORMATION: construct
US-09-206-866-41


  Query Match             100.0%;  Score 8;  DB 3;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
            ||||||||
Db       8 AACGTTCG 1


RESULT  19
US-09-206-866A-37/c
; Sequence 37, Application US/09206866A
; Patent No. 6268137
; GENERAL INFORMATION:
;  APPLICANT: SZYF, Moshe
;  APPLICANT:  BIGEY, Pascal
;  TITLE OF INVENTION: SPECIFIC INHIBITORS OF DNA METHYLTRANSFERASE
;  FILE REFERENCE: 106101.200
;  CURRENT APPLICATION NUMBER: US/09/206,866A
;  CURRENT FILING DATE:  1998-12-08
;  PRIOR APPLICATION NUMBER: US 08/653,954
;  PRIOR FILING DATE: 1996-05-22
;  PRIOR APPLICATION NUMBER: PCT/IB97/00879
;  PRIOR FILING DATE: 1997-05-22
;  PRIOR APPLICATION NUMBER: US 60/069,812
;  PRIOR FILING DATE: 1997-12-17
;  PRIOR APPLICATION NUMBER: US 09/194,284
;  PRIOR FILING DATE: 1998-11-23
```

```
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 37
;   LENGTH: 16
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(16)
;   OTHER INFORMATION: Nucleotides 1-16 contain c, t, a & g wherein
;   OTHER INFORMATION: c=cytidine; t=thymidine; a=adenosine; g=guanosine;
;   OTHER INFORMATION: m is a methyl group at the 5-position of
;   OTHER INFORMATION: nucleotides 1 & 5 of the cytosine porition of
;   OTHER INFORMATION: cytidine.
;   OTHER INFORMATION: Description of Artificial Sequence:synthetic
;   OTHER INFORMATION: construct
US-09-206-866A-37


  Query Match           100.0%;  Score 8;  DB 4;  Length 16;
  Best Local Similarity  100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       8 AACGTTCG 1


RESULT  20
US-09-206-866A-38/c
; Sequence 38, Application US/09206866A
; Patent No. 6268137
; GENERAL INFORMATION:
;  APPLICANT: SZYF, Moshe
;  APPLICANT:  BIGEY, Pascal
;  TITLE OF INVENTION: SPECIFIC INHIBITORS OF DNA METHYLTRANSFERASE
;  FILE REFERENCE: 106101.200
;  CURRENT APPLICATION NUMBER: US/09/206,866A
;  CURRENT FILING DATE:  1998-12-08
;  PRIOR APPLICATION NUMBER: US 08/653,954
;  PRIOR FILING DATE: 1996-05-22
;  PRIOR APPLICATION NUMBER: PCT/IB97/00879
;  PRIOR FILING DATE: 1997-05-22
;  PRIOR APPLICATION NUMBER: US 60/069,812
;  PRIOR FILING DATE: 1997-12-17
;  PRIOR APPLICATION NUMBER: US 09/194,284
;  PRIOR FILING DATE: 1998-11-23
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
;   LENGTH: 16
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(16)
;   OTHER INFORMATION: Nucleotides 1-16 contain c, t, a & g wherein
;   OTHER INFORMATION: c=cytidine; t=thymidine; a=adenosine; g=guanosine;
;   OTHER INFORMATION: m is a methyl group at the 5-position of
;   OTHER INFORMATION: nucleotides 1 & 5 of the cytosine portion of
;   OTHER INFORMATION: cytidine.
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(15)
;   OTHER INFORMATION: Nucleotide 15 is n wherein n = i and i = inosine.
;   OTHER INFORMATION: Description of Artificial Sequence:synthetic
;   OTHER INFORMATION: construct
US-09-206-866A-38


  Query Match           100.0%;  Score 8;  DB 4;  Length 16;
  Best Local Similarity  100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       8 AACGTTCG 1


RESULT  21
US-09-206-866A-39/c
; Sequence 39, Application US/09206866A
; Patent No. 6268137
; GENERAL INFORMATION:
;  APPLICANT: SZYF, Moshe
;  APPLICANT:  BIGEY, Pascal
;  TITLE OF INVENTION: SPECIFIC INHIBITORS OF DNA METHYLTRANSFERASE
;  FILE REFERENCE: 106101.200
;  CURRENT APPLICATION NUMBER: US/09/206,866A
;  CURRENT FILING DATE:  1998-12-08
;  PRIOR APPLICATION NUMBER: US 08/653,954
;  PRIOR FILING DATE: 1996-05-22
;  PRIOR APPLICATION NUMBER: PCT/IB97/00879
;  PRIOR FILING DATE: 1997-05-22
;  PRIOR APPLICATION NUMBER: US 60/069,812
;  PRIOR FILING DATE: 1997-12-17
;  PRIOR APPLICATION NUMBER: US 09/194,284
;  PRIOR FILING DATE: 1998-11-23
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
;   LENGTH: 16
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(16)
;   OTHER INFORMATION: Nucleotides 1-16 contain c, t, a & g wherein
;   OTHER INFORMATION: c=cytidine; t=thymidine; a=adenosine; g=guanosine;
;   OTHER INFORMATION: m is a methyl group at the 5-position of
;   OTHER INFORMATION: nucleotides 1 & 5 of the cytosine portion of
;   OTHER INFORMATION: cytidine.
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(15)
;   OTHER INFORMATION: Nucleotide 15 is n wherein n = u and u = uridine.
;   OTHER INFORMATION: Description of Artificial Sequence:synthetic
;   OTHER INFORMATION: construct
US-09-206-866A-39


  Query Match           100.0%;  Score 8;  DB 4;  Length 16;
  Best Local Similarity  100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       8 AACGTTCG 1


RESULT  22
US-09-206-866A-40/c
; Sequence 40, Application US/09206866A
; Patent No. 6268137
; GENERAL INFORMATION:
;  APPLICANT: SZYF, Moshe
;  APPLICANT:  BIGEY, Pascal
;  TITLE OF INVENTION: SPECIFIC INHIBITORS OF DNA METHYLTRANSFERASE
;  FILE REFERENCE: 106101.200
;  CURRENT APPLICATION NUMBER: US/09/206,866A
;  CURRENT FILING DATE:  1998-12-08
;  PRIOR APPLICATION NUMBER: US 08/653,954
;  PRIOR FILING DATE: 1996-05-22
;  PRIOR APPLICATION NUMBER: PCT/IB97/00879
;  PRIOR FILING DATE: 1997-05-22
;  PRIOR APPLICATION NUMBER: US 60/069,812
;  PRIOR FILING DATE: 1997-12-17
;  PRIOR APPLICATION NUMBER: US 09/194,284
```

```
;  PRIOR FILING DATE: 1998-11-23
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 40
;   LENGTH: 16
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(16)
;   OTHER INFORMATION: Nucleotides 1-16 contain c, t, a & g wherein
;   OTHER INFORMATION: c=cytidine; t=thymidine; a=adenosine; g=guanosine;
;   OTHER INFORMATION: m is a methyl group at the 5-position of
;   OTHER INFORMATION: nucleotides 1 & 5 of the cytosine portion of
;   OTHER INFORMATION: cytidine.
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(15)
;   OTHER INFORMATION: Nucleotide 15 is n wherein n = f and f =
;   OTHER INFORMATION: 5-fluorocytosine.
;   OTHER INFORMATION: Description of Artificial Sequence:synthetic
;   OTHER INFORMATION: construct
US-09-206-866A-40


  Query Match           100.0%;  Score 8;  DB 4;  Length 16;
  Best Local Similarity  100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
            ||||||||
Db       8 AACGTTCG 1


RESULT  23
US-09-206-866A-41/c
; Sequence 41, Application US/09206866A
; Patent No. 6268137
; GENERAL INFORMATION:
;  APPLICANT: SZYF, Moshe
;  APPLICANT:  BIGEY, Pascal
;  TITLE OF INVENTION: SPECIFIC INHIBITORS OF DNA METHYLTRANSFERASE
;  FILE REFERENCE: 106101.200
;  CURRENT APPLICATION NUMBER: US/09/206,866A
;  CURRENT FILING DATE:  1998-12-08
;  PRIOR APPLICATION NUMBER: US 08/653,954
;  PRIOR FILING DATE: 1996-05-22
;  PRIOR APPLICATION NUMBER: PCT/IB97/00879
;  PRIOR FILING DATE: 1997-05-22
;  PRIOR APPLICATION NUMBER: US 60/069,812
;  PRIOR FILING DATE: 1997-12-17
;  PRIOR APPLICATION NUMBER: US 09/194,284
;  PRIOR FILING DATE: 1998-11-23
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
;   LENGTH: 16
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(16)
;   OTHER INFORMATION: Nucleotides 1-16 contain c, t, a & g wherein
;   OTHER INFORMATION: c=cytidine; t=thymidine; a=adenosine; g=guanosine;
;   OTHER INFORMATION: m is a methyl group at the 5-position of
;   OTHER INFORMATION: nucleotides 1 & 5 of the cytosine portion of
;   OTHER INFORMATION: cytidine.
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(15)
;   OTHER INFORMATION: Nucleotide 15 is n wherein n = b and b = cytosine, inosine,
;   OTHER INFORMATION: uridine, 5-bromocytidine or 5-fluorouridine.
;   OTHER INFORMATION: Description of Artificial Sequence:synthetic
;   OTHER INFORMATION: construct
US-09-206-866A-41


  Query Match           100.0%;  Score 8;  DB 4;  Length 16;
  Best Local Similarity  100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
            ||||||||
Db       8 AACGTTCG 1


RESULT  24
US-09-206-866-20/c
; Sequence 20, Application US/09206866A
; Patent No. 6150108
; GENERAL INFORMATION:
;  APPLICANT: SZYF, Moshe
;  APPLICANT: BIGEY, Pascal
;  TITLE OF INVENTION: SPECIFIC INHIBITORS OF DNA METHYLTRANSFERASE
;  FILE REFERENCE: 106101.200
;  CURRENT APPLICATION NUMBER: US/09/206,866A
;  CURRENT FILING DATE: 1998-12-08
;  EARLIER APPLICATION NUMBER: US 08/653,954
;  EARLIER FILING DATE: 1996-05-22
;  EARLIER APPLICATION NUMBER: PCT/IB97/00879
;  EARLIER FILING DATE: 1997-05-22
;  EARLIER APPLICATION NUMBER: US 60/069,812
;  EARLIER FILING DATE: 1997-12-17
;  EARLIER APPLICATION NUMBER: US 09/194,284
;  EARLIER FILING DATE: 1998-11-23
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
;   LENGTH: 17
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(17)
;   OTHER INFORMATION: Nucleotides 1-17 contain c, t, a & g wherein
;   OTHER INFORMATION: c=cytidine; t=thymidine; a=adenosine; g=guanosine;
;   OTHER INFORMATION: m is a methyl group at the 5-position of
;   OTHER INFORMATION: nucleotides 1 & 5 of the cytosine portion of
;   OTHER INFORMATION: cytidine.
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:synthetic
;   OTHER INFORMATION: construct
US-09-206-866-20


  Query Match           100.0%;  Score 8;  DB 3;  Length 17;
  Best Local Similarity  100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
            ||||||||
Db       8 AACGTTCG 1


RESULT  25
US-09-206-866-21/c
; Sequence 21, Application US/09206866A
; Patent No. 6150108
; GENERAL INFORMATION:
;  APPLICANT: SZYF, Moshe
;  APPLICANT: BIGEY, Pascal
;  TITLE OF INVENTION: SPECIFIC INHIBITORS OF DNA METHYLTRANSFERASE
;  FILE REFERENCE: 106101.200
;  CURRENT APPLICATION NUMBER: US/09/206,866A
;  CURRENT FILING DATE: 1998-12-08
;  EARLIER APPLICATION NUMBER: US 08/653,954
;  EARLIER FILING DATE: 1996-05-22
;  EARLIER APPLICATION NUMBER: PCT/IB97/00879
```

```
;  EARLIER FILING DATE: 1997-05-22
;  EARLIER APPLICATION NUMBER: US 60/069,812
;  EARLIER FILING DATE: 1997-12-17
;  EARLIER APPLICATION NUMBER: US 09/194,284
;  EARLIER FILING DATE: 1998-11-23
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
;   LENGTH: 17
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(17)
;   OTHER INFORMATION: Nucleotides 1-17 contain c, t, a & g wherein
;   OTHER INFORMATION: c=cytidine; t=thymidine; a=adenosine; g=guanosine;
;   OTHER INFORMATION: m is a methyl group at the 5-position of
;   OTHER INFORMATION: nucleotides 1 & 5 of the cytosine portion of
;   OTHER INFORMATION: cytidine.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(16)
;   OTHER INFORMATION: Nucleotide 16 is n wherein n = i and i = inosine.
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:synthetic
;   OTHER INFORMATION: construct
US-09-206-866-21


  Query Match             100.0%;  Score 8;  DB 3;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
            ||||||||
Db       8 AACGTTCG 1


RESULT  26
US-09-206-866-22/c
; Sequence 22, Application US/09206866A
; Patent No. 6150108
; GENERAL INFORMATION:
;  APPLICANT: SZYF, Moshe
;  APPLICANT: BIGEY, Pascal
;  TITLE OF INVENTION: SPECIFIC INHIBITORS OF DNA METHYLTRANSFERASE
;  FILE REFERENCE: 106101.200
;  CURRENT APPLICATION NUMBER: US/09/206,866A
;  CURRENT FILING DATE: 1998-12-08
;  EARLIER APPLICATION NUMBER: US 08/653,954
;  EARLIER FILING DATE: 1996-05-22
;  EARLIER APPLICATION NUMBER: PCT/IB97/00879
;  EARLIER FILING DATE: 1997-05-22
;  EARLIER APPLICATION NUMBER: US 60/069,812
;  EARLIER FILING DATE: 1997-12-17
;  EARLIER APPLICATION NUMBER: US 09/194,284
;  EARLIER FILING DATE: 1998-11-23
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
;   LENGTH: 17
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(17)
;   OTHER INFORMATION: Nucleotides 1-17 contain c, t, a & g wherein
;   OTHER INFORMATION: c=cytidine; t=thymidine; a=adenosine; g=guanosine;
;   OTHER INFORMATION: m is a methyl group at the 5-position of
;   OTHER INFORMATION: nucleotides 1 & 5 of the cytosine portion of
;   OTHER INFORMATION: cytidine.
;   FEATURE:
;  , NAME/KEY: misc_feature
;   LOCATION: (1)..(16)
;   OTHER INFORMATION: Nucleotide 16 is n wherein n = u and u = uridine.
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:synthetic
;   OTHER INFORMATION: construct
US-09-206-866-22


  Query Match             100.0%;  Score 8;  DB 3;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
            ||||||||
Db       8 AACGTTCG 1


RESULT  27
US-09-206-866-23/c
; Sequence 23, Application US/09206866A
; Patent No. 6150108
; GENERAL INFORMATION:
;  APPLICANT: SZYF, Moshe
;  APPLICANT: BIGEY, Pascal
;  TITLE OF INVENTION: SPECIFIC INHIBITORS OF DNA METHYLTRANSFERASE
;  FILE REFERENCE: 106101.200
;  CURRENT APPLICATION NUMBER: US/09/206,866A
;  CURRENT FILING DATE: 1998-12-08
;  EARLIER APPLICATION NUMBER: US 08/653,954
;  EARLIER FILING DATE: 1996-05-22
;  EARLIER APPLICATION NUMBER: PCT/IB97/00879
;  EARLIER FILING DATE: 1997-05-22
;  EARLIER APPLICATION NUMBER: US 60/069,812
;  EARLIER FILING DATE: 1997-12-17
;  EARLIER APPLICATION NUMBER: US 09/194,284
;  EARLIER FILING DATE: 1998-11-23
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
;   LENGTH: 17
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(17)
;   OTHER INFORMATION: Nucleotides 1-17 contain c, t, a & g wherein
;   OTHER INFORMATION: c=cytidine; t=thymidine; a=adenosine; g=guanosine;
;   OTHER INFORMATION: m is a methyl group at the 5-position of
;   OTHER INFORMATION: nucleotides 1 & 5 of the cytosine portion of
;   OTHER INFORMATION: cytidine.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(16)
;   OTHER INFORMATION: Nucleotides 12 & 16 are n wherein n = f and f =
;   OTHER INFORMATION: 5-fluorocytosine.
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:synthetic
;   OTHER INFORMATION: construct
US-09-206-866-23


  Query Match             100.0%;  Score 8;  DB 3;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
            ||||||||
Db       8 AACGTTCG 1


RESULT  28
US-09-206-866-24/c
; Sequence 24, Application US/09206866A
```

```
; Patent No. 6150108
; GENERAL INFORMATION:
;  APPLICANT: SZYF, Moshe
;  APPLICANT: BIGEY, Pascal
;  TITLE OF INVENTION: SPECIFIC INHIBITORS OF DNA METHYLTRANSFERASE
;  FILE REFERENCE: 106101.200
;  CURRENT APPLICATION NUMBER: US/09/206,866A
;  CURRENT FILING DATE: 1998-12-08
;  EARLIER APPLICATION NUMBER: US 08/653,954
;  EARLIER FILING DATE: 1996-05-22
;  EARLIER APPLICATION NUMBER: PCT/IB97/00879
;  EARLIER FILING DATE: 1997-05-22
;  EARLIER APPLICATION NUMBER: US 60/069,812
;  EARLIER FILING DATE: 1997-12-17
;  EARLIER APPLICATION NUMBER: US 09/194,284
;  EARLIER FILING DATE: 1998-11-23
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
;   LENGTH: 17
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(17)
;   OTHER INFORMATION: Nucleotides 1-17 contain c, t, a & g wherein
;   OTHER INFORMATION: c=cytidine; t=thymidine; a=adenosine; g=guanosine;
;   OTHER INFORMATION: m is a methyl group at the 5-position of
;   OTHER INFORMATION: nucleotides 1 & 5 of the cytosine portion of
;   OTHER INFORMATION: cytidine.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(16)
;   OTHER INFORMATION: Nucleotides 12 & 16 are n wherein n = b and b =
;   OTHER INFORMATION: cytosine, inosine, uridine, 5-bromocytidine or
;   OTHER INFORMATION: 5-fluorouridine.
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:synthetic
;   OTHER INFORMATION: construct
US-09-206-866-24


  Query Match           100.0%;  Score 8;  DB 3;  Length 17;
  Best Local Similarity  100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
            ||||||||
Db       8 AACGTTCG 1


RESULT  29
US-09-206-866A-20/c
; Sequence 20, Application US/09206866A
; Patent No. 6268137
; GENERAL INFORMATION:
;  APPLICANT: SZYF, Moshe
;  APPLICANT:  BIGEY, Pascal
;  TITLE OF INVENTION: SPECIFIC INHIBITORS OF DNA METHYLTRANSFERASE
;  FILE REFERENCE: 106101.200
;  CURRENT APPLICATION NUMBER: US/09/206,866A
;  CURRENT FILING DATE:  1998-12-08
;  PRIOR APPLICATION NUMBER: US 08/653,954
;  PRIOR FILING DATE: 1996-05-22
;  PRIOR APPLICATION NUMBER: PCT/IB97/00879
;  PRIOR FILING DATE: 1997-05-22
;  PRIOR APPLICATION NUMBER: US 60/069,812
;  PRIOR FILING DATE: 1997-12-17
;  PRIOR APPLICATION NUMBER: US 09/194,284
;  PRIOR FILING DATE: 1998-11-23
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
;   LENGTH: 17
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(17)
;   OTHER INFORMATION: Nucleotides 1-17 contain c, t, a & g wherein
;   OTHER INFORMATION: c=cytidine; t=thymidine; a=adenosine; g=guanosine;
;   OTHER INFORMATION: m is a methyl group at the 5-position of
;   OTHER INFORMATION: nucleotides 1 & 5 of the cytosine portion of
;   OTHER INFORMATION: cytidine.
;   OTHER INFORMATION: Description of Artificial Sequence:synthetic
;   OTHER INFORMATION: construct
US-09-206-866A-20


  Query Match           100.0%;  Score 8;  DB 4;  Length 17;
  Best Local Similarity  100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
            ||||||||
Db       8 AACGTTCG 1


RESULT  30
US-09-206-866A-21/c
; Sequence 21, Application US/09206866A
; Patent No. 6268137
; GENERAL INFORMATION:
;  APPLICANT: SZYF, Moshe
;  APPLICANT:  BIGEY, Pascal
;  TITLE OF INVENTION: SPECIFIC INHIBITORS OF DNA METHYLTRANSFERASE
;  FILE REFERENCE: 106101.200
;  CURRENT APPLICATION NUMBER: US/09/206,866A
;  CURRENT FILING DATE:  1998-12-08
;  PRIOR APPLICATION NUMBER: US 08/653,954
;  PRIOR FILING DATE: 1996-05-22
;  PRIOR APPLICATION NUMBER: PCT/IB97/00879
;  PRIOR FILING DATE: 1997-05-22
;  PRIOR APPLICATION NUMBER: US 60/069,812
;  PRIOR FILING DATE: 1997-12-17
;  PRIOR APPLICATION NUMBER: US 09/194,284
;  PRIOR FILING DATE: 1998-11-23
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
;   LENGTH: 17
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(17)
;   OTHER INFORMATION: Nucleotides 1-17 contain c, t, a & g wherein
;   OTHER INFORMATION: c=cytidine; t=thymidine; a=adenosine; g=guanosine;
;   OTHER INFORMATION: m is a methyl group at the 5-position of
;   OTHER INFORMATION: nucleotides 1 & 5 of the cytosine portion of
;   OTHER INFORMATION: cytidine.
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(16)
;   OTHER INFORMATION: Nucleotide 16 is n wherein n = i and i = inosine.
;   OTHER INFORMATION: Description of Artificial Sequence:synthetic
;   OTHER INFORMATION: construct
US-09-206-866A-21


  Query Match           100.0%;  Score 8;  DB 4;  Length 17;
  Best Local Similarity  100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
            ||||||||
Db       8 AACGTTCG 1
```

```
RESULT 31
US-09-206-866A-22/c
; Sequence 22, Application US/09206866A
; Patent No. 6268137
; GENERAL INFORMATION:
;  APPLICANT: SZYF, Moshe
;  APPLICANT:  BIGEY, Pascal
;  TITLE OF INVENTION: SPECIFIC INHIBITORS OF DNA METHYLTRANSFERASE
;  FILE REFERENCE: 106101.200
;  CURRENT APPLICATION NUMBER: US/09/206,866A
;  CURRENT FILING DATE:  1998-12-08
;  PRIOR APPLICATION NUMBER: US 08/653,954
;  PRIOR FILING DATE: 1996-05-22
;  PRIOR APPLICATION NUMBER: PCT/IB97/00879
;  PRIOR FILING DATE: 1997-05-22
;  PRIOR APPLICATION NUMBER: US 60/069,812
;  PRIOR FILING DATE: 1997-12-17
;  PRIOR APPLICATION NUMBER: US 09/194,284
;  PRIOR FILING DATE: 1998-11-23
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
;   LENGTH: 17
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(17)
;   OTHER INFORMATION: Nucleotides 1-17 contain c, t, a & g wherein
;   OTHER INFORMATION: c=cytidine; t=thymidine; a=adenosine; g=guanosine;
;   OTHER INFORMATION: m is a methyl group at the 5-position of
;   OTHER INFORMATION: nucleotides 1 & 5 of the cytosine portion of
;   OTHER INFORMATION: cytidine.
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(16)
;   OTHER INFORMATION: Nucleotide 16 is n wherein n = u and u = uridine.
;   OTHER INFORMATION: Description of Artificial Sequence:synthetic
;   OTHER INFORMATION: construct
US-09-206-866A-22


  Query Match           100.0%;  Score 8;  DB 4;  Length 17;
  Best Local Similarity  100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
            ||||||||
Db       8 AACGTTCG 1


RESULT 32
US-09-206-866A-23/c
; Sequence 23, Application US/09206866A
; Patent No. 6268137
; GENERAL INFORMATION:
;  APPLICANT: SZYF, Moshe
;  APPLICANT:  BIGEY, Pascal
;  TITLE OF INVENTION: SPECIFIC INHIBITORS OF DNA METHYLTRANSFERASE
;  FILE REFERENCE: 106101.200
;  CURRENT APPLICATION NUMBER: US/09/206,866A
;  CURRENT FILING DATE:  1998-12-08
;  PRIOR APPLICATION NUMBER: US 08/653,954
;  PRIOR FILING DATE: 1996-05-22
;  PRIOR APPLICATION NUMBER: PCT/IB97/00879
;  PRIOR FILING DATE: 1997-05-22
;  PRIOR APPLICATION NUMBER: US 60/069,812
;  PRIOR FILING DATE: 1997-12-17
;  PRIOR APPLICATION NUMBER: US 09/194,284
;  PRIOR FILING DATE: 1998-11-23
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
;   LENGTH: 17
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(17)
;   OTHER INFORMATION: Nucleotides 1-17 contain c, t, a & g wherein
;   OTHER INFORMATION: c=cytidine; t=thymidine; a=adenosine; g=guanosine;
;   OTHER INFORMATION: m is a methyl group at the 5-position of
;   OTHER INFORMATION: nucleotides 1 & 5 of the cytosine portion of
;   OTHER INFORMATION: cytidine.
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(16)
;   OTHER INFORMATION: Nucleotides 12 & 16 are n wherein n = f and f =
;   OTHER INFORMATION: 5-fluorocytosine.
;   OTHER INFORMATION: Description of Artificial Sequence:synthetic
;   OTHER INFORMATION: construct
US-09-206-866A-23


  Query Match           100.0%;  Score 8;  DB 4;  Length 17;
  Best Local Similarity  100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
            ||||||||
Db       8 AACGTTCG 1


RESULT 33
US-09-206-866A-24/c
; Sequence 24, Application US/09206866A
; Patent No. 6268137
; GENERAL INFORMATION:
;  APPLICANT: SZYF, Moshe
;  APPLICANT:  BIGEY, Pascal
;  TITLE OF INVENTION: SPECIFIC INHIBITORS OF DNA METHYLTRANSFERASE
;  FILE REFERENCE: 106101.200
;  CURRENT APPLICATION NUMBER: US/09/206,866A
;  CURRENT FILING DATE:  1998-12-08
;  PRIOR APPLICATION NUMBER: US 08/653,954
;  PRIOR FILING DATE: 1996-05-22
;  PRIOR APPLICATION NUMBER: PCT/IB97/00879
;  PRIOR FILING DATE: 1997-05-22
;  PRIOR APPLICATION NUMBER: US 60/069,812
;  PRIOR FILING DATE: 1997-12-17
;  PRIOR APPLICATION NUMBER: US 09/194,284
;  PRIOR FILING DATE: 1998-11-23
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
;   LENGTH: 17
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(17)
;   OTHER INFORMATION: Nucleotides 1-17 contain c, t, a & g wherein
;   OTHER INFORMATION: c=cytidine; t=thymidine; a=adenosine; g=guanosine;
;   OTHER INFORMATION: m is a methyl group at the 5-position of
;   OTHER INFORMATION: nucleotides 1 & 5 of the cytosine portion of
;   OTHER INFORMATION: cytidine.
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(16)
;   OTHER INFORMATION: Nucleotides 12 & 16 are n wherein n = b and b =
;   OTHER INFORMATION: cytosine, inosine, uridine, 5-bromocytidine or
;   OTHER INFORMATION: 5-fluorouridine.
;   OTHER INFORMATION: Description of Artificial Sequence:synthetic
;   OTHER INFORMATION: construct
US-09-206-866A-24
```

```
  Query Match             100.0%;  Score 8;  DB 4;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       8 AACGTTCG 1


RESULT  34
US-08-255-892-37
; Sequence 37, Application US/08255892
; Patent No. 5695926
;  GENERAL INFORMATION:
;    APPLICANT:  CROS, PHILIPPE
;    APPLICANT:  ALLIBERT, PATRICE
;    APPLICANT:  MALLET, FRANCOIS
;    APPLICANT:  MABILAT, CLAUDE
;    APPLICANT:  MANDRAND, BERNARD
;    TITLE OF INVENTION:  PROCEDURE FOR DETECTION OF A NUCLEOTIDE
;    TITLE OF INVENTION:  SEQUENCE BY IMPLEMENTING THE SANDWICH HYBRIDIZATION
;    TITLE OF INVENTION:  TECHNIQUE
;    NUMBER OF SEQUENCES:  113
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CUSHMAN, DARBY & CUSHMAN
;      STREET:  1100 NEW YORK AVENUE, N.W.
;      CITY:  WASHINGTON
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/255,892
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/834,543
;      FILING DATE:  11-FEB-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DEAVER, DONALD B.
;      REGISTRATION NUMBER:  23,048
;      REFERENCE/DOCKET NUMBER:  1032/94109
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-861-3000
;      TELEFAX:  202-822-0944
;      TELEX:  6714627 CUSH
;  INFORMATION FOR SEQ ID NO:  37:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-255-892-37


  Query Match             100.0%;  Score 8;  DB 1;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       5 AACGTTCG 12


RESULT  35
US-08-506-864A-11/c
; Sequence 11, Application US/08506864A
; Patent No. 5834245
;  GENERAL INFORMATION:
;    APPLICANT:  NAKAMURA, YUSUKE
;    APPLICANT:  : FUJIWARA, YOSHIYUKI
;    TITLE OF INVENTION:  PRLTS PROTEINS AND DNA'S
;    TITLE OF INVENTION:  ENCODING THE SAME
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FLYNN, THIEL, BOUTELL & TANIS, P.C.
;      STREET:  2026 Rambling Road
;      CITY:  Kalamazoo
;      STATE:  Michigan
;      COUNTRY:  USA
;      ZIP:  49008-1699
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inches, 1.44 Mb storage
;      COMPUTER:  IBM PC/XT/AT Compatible
;      OPERATING SYSTEM:  MS-DOS 5.0
;      SOFTWARE:  WordPerfect 5.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/506,864A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP6-178131
;      FILING DATE:  29-JULY-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Terryence F. Chapman
;      REGISTRATION NUMBER:  32549
;      REFERENCE/DOCKET NUMBER:  Furuya Case 1334
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (616) 381-1156
;      TELEFAX:  (616) 381-5465
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid (synthetic DNA)
US-08-506-864A-11


  Query Match             100.0%;  Score 8;  DB 2;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      17 AACGTTCG 10


RESULT  36
US-08-851-968-11/c
; Sequence 11, Application US/08851968
; Patent No. 5935786
;  GENERAL INFORMATION:
;    APPLICANT:  NAKAMURA, YUSUKE
;    APPLICANT:  : FUJIWARA, YOSHIYUKI
;    TITLE OF INVENTION:  PRLTS PROTEINS AND DNA'S
;    TITLE OF INVENTION:  ENCODING THE SAME
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FLYNN, THIEL, BOUTELL & TANIS, P.C.
;      STREET:  2026 Rambling Road
;      CITY:  Kalamazoo
;      STATE:  Michigan
;      COUNTRY:  USA
;      ZIP:  49008-1699
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inches, 1.44 Mb storage
;      COMPUTER:  IBM PC/XT/AT Compatible
;      OPERATING SYSTEM:  MS-DOS 5.0
;      SOFTWARE:  WordPerfect 5.0
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/851,968
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/506,864
;      FILING DATE:
;      APPLICATION NUMBER:  JP6-178131
;      FILING DATE:  29-JULY-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Terryence F. Chapman
;      REGISTRATION NUMBER:  32549
;      REFERENCE/DOCKET NUMBER:  Furuya Case 1334
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (616) 381-1156
;      TELEFAX:  (616) 381-5465
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid (synthetic DNA)
US-08-851-968-11


  Query Match           100.0%;  Score 8;  DB 2;  Length 20;
  Best Local Similarity  100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
            ||||||||
Db      17 AACGTTCG 10


RESULT  37
US-09-286-098-11/c
; Sequence 11, Application US/09286098
; Patent No. 6218371
; GENERAL INFORMATION:
;  APPLICANT: Krieg, Arthur M.
;  APPLICANT: Weiner, George
;  TITLE OF INVENTION: Methods and Products for Stimulating the
;  TITLE OF INVENTION: Immune System Using Immunotherapeutic Oligonucleotides and
;  TITLE OF INVENTION: Cytokines
;  FILE REFERENCE: C1039/7026/HCL
;  CURRENT APPLICATION NUMBER: US/09/286,098
;  CURRENT FILING DATE: 1999-04-02
;  EARLIER APPLICATION NUMBER: US 60/080,729
;  EARLIER FILING DATE: 1998-04-03
;  NUMBER OF SEQ ID NOS: 105
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 11
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic Sequence
US-09-286-098-11


  Query Match           100.0%;  Score 8;  DB 4;  Length 20;
  Best Local Similarity  100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
            ||||||||
Db      17 AACGTTCG 10


RESULT  38
US-08-882-704A-18
; Sequence 18, Application US/08882704A
; Patent No. 5879906
; GENERAL INFORMATION:
;   APPLICANT:  Jefferson, Richard A.
;   APPLICANT:  Wilson, Katherine J.
;   APPLICANT:  Leader, Michael
;   TITLE OF INVENTION:  GLUCURONIDE REPRESSORS AND USES THEREOF
;   NUMBER OF SEQUENCES:  19
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  SEED and BERRY LLP
;     STREET:  6300 Columbia Center, 701 Fifth Avenue
;     CITY:  Seattle
;     STATE:  Washington
;     COUNTRY:  USA
;     ZIP:  98104-7092
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/882,704A
;     FILING DATE:  25-JUN-1997
;     CLASSIFICATION:  435
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  No. 5879906tenburg Ph.D., Carol
;     REGISTRATION NUMBER:  39,317
;     REFERENCE/DOCKET NUMBER:  190106.404
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  (206) 622-4900
;     TELEFAX:  (206) 682-6031
; INFORMATION FOR SEQ ID NO:  18:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  22 base pairs
;     TYPE:  nucleic acid
;     STRANDEDNESS:  single
;     TOPOLOGY:  linear
US-08-882-704A-18


  Query Match           100.0%;  Score 8;  DB 2;  Length 22;
  Best Local Similarity  100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
            ||||||||
Db      11 AACGTTCG 18


RESULT  39
US-08-882-704A-18/c
; Sequence 18, Application US/08882704A
; Patent No. 5879906
; GENERAL INFORMATION:
;   APPLICANT:  Jefferson, Richard A.
;   APPLICANT:  Wilson, Katherine J.
;   APPLICANT:  Leader, Michael
;   TITLE OF INVENTION:  GLUCURONIDE REPRESSORS AND USES THEREOF
;   NUMBER OF SEQUENCES:  19
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  SEED and BERRY LLP
;     STREET:  6300 Columbia Center, 701 Fifth Avenue
;     CITY:  Seattle
;     STATE:  Washington
;     COUNTRY:  USA
;     ZIP:  98104-7092
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/882,704A
;     FILING DATE:  25-JUN-1997
```

```
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5879906tenburg Ph.D., Carol
;      REGISTRATION NUMBER:  39,317
;      REFERENCE/DOCKET NUMBER:  190106.404
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  22 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-882-704A-18


  Query Match             100.0%;  Score 8;  DB 2;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      16 AACGTTCG 9


RESULT  40
US-08-064-271-9/c
; Sequence 9, Application US/08064271
; Patent No. 5543297
;  GENERAL INFORMATION:
;    APPLICANT:  Kennedy, Brian P.
;    APPLICANT:  Cromlish, Wanda A.
;    APPLICANT:  Mancini, Joseph A.
;    APPLICANT:  O'Neil, Gary
;    APPLICANT:  Vickers, Philip J.
;    APPLICANT:  Wong, Elizabeth
;    TITLE OF INVENTION:  HUMAN CYCLOOXYGENASE-2 cDNA AND
;    TITLE OF INVENTION:  ASSAY FOR EVALUATING CYCLOOXYGENASE ACTIVITY
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merck & Co., Inc.
;      STREET:  126 Lincoln Avenue
;      CITY:  Rahway
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 in, 1.4kb
;      COMPUTER:  Apple Macintosh
;      OPERATING SYSTEM:  System 7
;      SOFTWARE:  Microsoft Word 5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/064,271
;      FILING DATE:  19930506
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Panzer, Curtis C.
;      REGISTRATION NUMBER:  33,752
;      REFERENCE/DOCKET NUMBER:  18906IA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (908)594-3199
;      TELEFAX:  (908)594-4720
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  24 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-064-271-9


  Query Match             100.0%;  Score 8;  DB 1;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      23 AACGTTCG 16


RESULT  41
US-08-930-589A-9/c
; Sequence 9, Application US/08930589A
; Patent No. 6107087
;  GENERAL INFORMATION:
;    APPLICANT:  MERCK FROSST CANADA & CO.
;    APPLICANT:  O'NEILL, GARY P.
;    APPLICANT:  MANCINI, JOSEPH A.
;    TITLE OF INVENTION:  HIGH LEVEL EXPRESSION OF HUMAN
;    TITLE OF INVENTION:  CYCLOOXYGENASE-2
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merck & Co., Inc.
;      STREET:  P.O. Box 2000, 126 E. Lincoln Ave.
;      CITY:  Rahway
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065-0900
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/930,589A
;      FILING DATE:  28-JUN-1998
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coppola, Joseph A
;      REGISTRATION NUMBER:  38,413
;      REFERENCE/DOCKET NUMBER:  19029PC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-6734
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  24 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-930-589A-9


  Query Match             100.0%;  Score 8;  DB 3;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      23 AACGTTCG 16


RESULT  42
US-09-129-686-16/c
; Sequence 16, Application US/09129686A
; Patent No. 6264940
; GENERAL INFORMATION:
;  APPLICANT: Gromeier PhD, Matthias
;  APPLICANT: Wimmer Prof, Eckard
```

```
;  TITLE OF INVENTION: Recombinant Poliovirus For The Treatment of Cancer
;  FILE REFERENCE: Recomb Poliovirus for Cancer Treatment
;  CURRENT APPLICATION NUMBER: US/09/129,686A
;  CURRENT FILING DATE: 1998-08-05
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
;   LENGTH: 26
;   TYPE: DNA
;   ORGANISM: Human rhinovirus 2
US-09-129-686-16


  Query Match           100.0%;  Score 8;  DB 4;  Length 26;
  Best Local Similarity  100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 AACGTTCG 8
          ||||||||
Db     12 AACGTTCG 5


RESULT  43
US-08-081-070-8
; Sequence 8, Application US/08081070
; Patent No. 5306619
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    TITLE OF INVENTION:  Screening Assay for the Detection of
;    TITLE OF INVENTION:  DNA-Binding Molecules
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dehlinger & Swiss
;      STREET:  P.O. Box 60850
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/081,070
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/723,618
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  4600-0085
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 323-8302
;      TELEFAX:  (415) 323-8306
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  30 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  UL9 polyA TEST SEQ. / UL9 ASSAY SEQ.
US-08-081-070-8


  Query Match           100.0%;  Score 8;  DB 1;  Length 30;
  Best Local Similarity  100.0%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 AACGTTCG 8
          ||||||||
Db      8 AACGTTCG 15


RESULT  44
US-08-171-389-608
; Sequence 608, Application US/08171389
; Patent No. 5578444
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  641
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/171,389
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,070
;      FILING DATE:  22-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  4600-0175/G19P3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  608:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  30 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  UL9 polyA TEST SEQ. / UL9 ASSAY
;      INDIVIDUAL ISOLATE:  SEQ.
US-08-171-389-608


  Query Match           100.0%;  Score 8;  DB 1;  Length 30;
```

Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 AACGTTCG 8
            ||||||||
Db       8 AACGTTCG 15
```

RESULT 45
PCT-US93-12388-608
; Sequence 608, Application PC/TUS9312388
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Sequence-Directed DNA Binding
; TITLE OF INVENTION: Molecules, Compositions and Methods
; NUMBER OF SEQUENCES: 641
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genelabs Technologies, Inc.
; STREET: 505 Penobscot Drive
; CITY: Redwood City
; STATE: CA
; COUNTRY: USA
; ZIP: 94063
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/12388
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/123,936
; FILING DATE: 17-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/996,783
; FILING DATE: 23-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0175.41/G19PCT2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 608:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: UL9 polyA TEST SEQ. / UL9 ASSAY
; INDIVIDUAL ISOLATE: SEQ.
PCT-US93-12388-608

Query Match 100.0%; Score 8; DB 5; Length 30;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 AACGTTCG 8
            ||||||||
Db       8 AACGTTCG 15
```

Search completed: November 29, 2001, 14:48:17
Job time: 3590 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: November 29, 2001, 12:09:07 ; Search time 1878.42 Seconds
(without alignments)
45.765 Million cell updates/sec

Title: FRAG1
Perfect score: 8
Sequence: 1 AACGTTCG 8

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 260912

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_htc:*
10: gb_est1:*
11: gb_est2:*
12: gb_htc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	8	100.0	38	13	TA335E03Q	AL492118 T. brucei
2	8	100.0	50	10	AU104223	AU104223 AU104223
c 3	8	100.0	54	13	AZ300935	AZ300935 EP(2)2185
4	8	100.0	55	13	AZ785311	AZ785311 2M0029E07
c 5	8	100.0	57	13	TA93B08P	AL458792 T. brucei
6	8	100.0	62	10	AU008219	AU008219 AU008219
7	8	100.0	62	10	AU008222	AU008222 AU008222
8	8	100.0	62	10	AU008233	AU008233 AU008233
9	8	100.0	62	10	AU008237	AU008237 AU008237
10	8	100.0	64	10	BE638333	BE638333 SWOvMfCAR
11	8	100.0	64	11	BI097404	BI097404 SWOv3MCAM
c 12	8	100.0	67	10	AA617006	AA617006 vk51a11.r
c 13	8	100.0	69	13	AQ025258	AQ025258 EP(3)3076
c 14	8	100.0	71	10	BE024070	BE024070 sm96c11.y
15	8	100.0	73	10	AW499129	AW499129 SWOvAFCAP
16	8	100.0	74	10	AA404533	AA404533 zw37h02.s
17	8	100.0	76	10	BE027432	BE027432 EtESTea43
c 18	8	100.0	80	13	TA389C08P	AL498967 T. brucei
19	8	100.0	81	10	AI903642	AI903642 QV-BT032-
20	8	100.0	81	10	BE027387	BE027387 EtESTea43
21	8	100.0	85	10	AA629864	AA629864 ad48h11.s
c 22	8	100.0	85	10	AA629864	AA629864 ad48h11.s
23	8	100.0	85	10	AA670169	AA670169 ab65d05.s
c 24	8	100.0	85	10	AA670169	AA670169 ab65d05.s
c 25	8	100.0	86	13	TA245G05Q	AL482195 T. brucei
c 26	8	100.0	88	10	AI289175	AI289175 qn25f09.x
c 27	8	100.0	88	10	AA626216	AA626216 zv88a05.s
c 28	8	100.0	94	10	AJ239919	AJ239919 AJ239919
29	8	100.0	94	10	BE576515	BE576515 dc40g03.y
c 30	8	100.0	94	10	BE576515	BE576515 dc40g03.y
c 31	8	100.0	98	10	AU013893	AU013893 AU013893
32	8	100.0	99	10	AI105877	AI105877 ab01f04_t
33	8	100.0	100	10	AW609278	AW609278 MR3-ST019
34	7	87.5	22	10	AA991491	AA991491 os91h12.s
c 35	7	87.5	22	13	AZ430042	AZ430042 1M0214O13
36	7	87.5	25	13	TA114E04P	AL462601 T. brucei
37	7	87.5	26	13	AZ823311	AZ823311 2M0097H17
c 38	7	87.5	26	13	TA69D08Q	AL458491 T. brucei
39	7	87.5	27	13	AZ974368	AZ974368 2M0248J21
40	7	87.5	28	10	AA027602	AA027602 mi12d08.r
41	7	87.5	28	10	AI441029	AI441029 sa58e02.y
c 42	7	87.5	31	10	AA901420	AA901420 SWOv3MCAM
43	7	87.5	31	10	AA910190	AA910190 oj29a03.s
c 44	7	87.5	31	10	AA910190	AA910190 oj29a03.s
45	7	87.5	34	10	AI002314	AI002314 or73c09.s

ALIGNMENTS

RESULT 1
TA335E03Q/c
LOCUS TA335E03Q 38 bp DNA GSS 13-DEC-2000
DEFINITION T. brucei sheared genomic DNA clone 335e03, reverse sequence, genomic survey sequence.
ACCESSION AL492118
VERSION AL492118.1 GI:11868418
KEYWORDS GSS.
SOURCE Trypanosoma brucei.
ORGANISM Trypanosoma brucei
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
REFERENCE 1 (bases 1 to 38)
AUTHORS Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R., Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L., Melville,S.E., Rajandream,M.A. and Barrell,B.G.
TITLE Direct Submission
JOURNAL Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and nhl@sanger.ac.uk
COMMENT Constructed at the Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from a cloned population of Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared to give a tight size distribution (4 kb). The v + i method used for the library construction is described in detail in Smith, H. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999).
Email: nelsayed@tigr.org
Details of T. brucei sequencing at the Sanger Centre are available at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES Location/Qualifiers
source 1. .38

```
                     /organism="Trypanosoma brucei"
                     /strain="TREU927"
                     /db_xref="taxon:5691"
                     /clone="335e03"
BASE COUNT         7 a       9 c       9 g      13 t
ORIGIN


 Query Match             100.0%;  Score 8;  DB 13;  Length 38;
 Best Local Similarity   100.0%;  Pred. No. 3e+04;
 Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      16 AACGTTCG 9


RESULT   2
AU104223
LOCUS       AU104223        50 bp    mRNA            EST       05-APR-2001
DEFINITION  AU104223 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HEP21349, mRNA sequence.
ACCESSION   AU104223
VERSION     AU104223.1  GI:13553744
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
            ,K., Suyama,A. and Sugano,S.
  TITLE     Fine Structural analysis of transcription start sites of human
            mRNAs using full-length enriched and 5'-end enriched cDNA libraries
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HEP21349"
                     /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT         7 a      16 c      14 g      13 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 50;
  Best Local Similarity   100.0%;  Pred. No. 3.1e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      25 AACGTTCG 32


RESULT   3
AZ300935/c
LOCUS       AZ300935        54 bp    DNA             GSS       23-AUG-2000
DEFINITION  EP(2)2185 Drosophila melanogaster EP line Drosophila melanogaster
            genomic Both 5' and 3' ends of P element, DNA sequence.
ACCESSION   AZ300935
VERSION     AZ300935.1  GI:9650436
KEYWORDS    GSS.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 54)
  AUTHORS   Liao,G.-C., Rehm,E.J. and Rubin,G.M.
  TITLE     Insertion site preferences of the P transposable element in
            Drosophila melanogaster
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3347-3351 (2000)
  MEDLINE   20202638
COMMENT     Contact: Gerald Rubin
            Berkeley Drosophila Genome Project
            University of California, Berkeley
            LSA Building, Berkeley, CA 94720-3200, USA
            Fax: 5106439947
            Email: gerry@fruitfly.berkeley.edu
            Sequence recovery method was inverse PCR.

            Sequence orientation is forward strand relative to 5' end of P
            element

            The P element insertion position is base 1 in the 54 bases. This
            insertion position refers to the first base of the 8 base target
            recognition sequence.
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .54
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone_lib="Drosophila melanogaster EP line"
                     /note="Inverse PCR was performed on Drosophila
                     melanogaster strains each of which contains a single EP
                     transposable element insertion. (The generation of these
                     insertion strains is described in Rorth P, Szabo K, Bailey
                     A, Laverty T, Rehm J, Rubin GM, Weigmann K, Milan M, Benes
                     V, Ansorge W, Cohen SM. 1998. Systematic gain-of-function
                     genetics in Drosophila. Development 6:1049-1057.) The
                     resultant fragment for each strain was directly sequenced
                     to determine the genomic sequence at the site of
                     insertion. Details of the protocols used can be found at
                     http://fruitfly.berkeley.edu/p_disrupt/inverse_pcr.html."
BASE COUNT        10 a      12 c      20 g      12 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 13;  Length 54;
  Best Local Similarity   100.0%;  Pred. No. 3.2e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      19 AACGTTCG 12


RESULT   4
AZ785311
LOCUS       AZ785311        55 bp    DNA             GSS       16-FEB-2001
DEFINITION  2M0029E07F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0029E07 F, DNA sequence.
ACCESSION   AZ785311
VERSION     AZ785311.1  GI:12921925
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 55)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
```

```
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA .
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0029  row: E  column: 07
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 55.
FEATURES             Location/Qualifiers
     source          1. .55
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0029E07"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       18 a      9 c      9 g     19 t
ORIGIN

  Query Match             100.0%;  Score 8;  DB 13;  Length 55;
  Best Local Similarity   100.0%;  Pred. No. 3.2e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      35 AACGTTCG 42


RESULT   5
TA93B08P/c
LOCUS       TA93B08P       57 bp    DNA          GSS       13-DEC-2000
DEFINITION  T. brucei sheared genomic DNA clone 93b08, forward sequence,
            genomic survey sequence.
ACCESSION   AL458792
VERSION     AL458792.1  GI:11861264
KEYWORDS    GSS.
SOURCE      Trypanosoma brucei.
  ORGANISM  Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 57)
  AUTHORS   Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
            Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
            Melville,S.E., Rajandream,M.A. and Barrell,B.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
            project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
            Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
            nh1@sanger.ac.uk
COMMENT     Constructed at the Institute for Genomic Research (TIGR),
            Rockville, MD. Genomic DNA isolated from a cloned population of
            Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
            to give a tight size distribution (
            4 kb).  The v + i method used for the library construction is
            described in detail in Smith, H. and Venter, J.C. (Making small
            insert libraries for whole genome shotgun sequencing projects.  In
            Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
            Barrell, Oxford University Press, 1999).
            Email: nelsayed@tigr.org
            Details of T. brucei sequencing at the Sanger Centre are available
            at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES             Location/Qualifiers
     source          1. .57
                     /organism="Trypanosoma brucei"
                     /strain="TREU927"
                     /db_xref="taxon:5691"
                     /clone="93b08"
BASE COUNT       12 a      14 c      18 g      13 t
ORIGIN

  Query Match             100.0%;  Score 8;  DB 13;  Length 57;
  Best Local Similarity   100.0%;  Pred. No. 3.2e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      48 AACGTTCG 41


RESULT   6
AU008219
LOCUS       AU008219       62 bp    mRNA            EST       31-JUL-1998
DEFINITION  AU008219 Schizosaccharomyces pombe late log phase cDNA
            Schizosaccharomyces pombe cDNA clone spc03066, mRNA sequence.
ACCESSION   AU008219
VERSION     AU008219.1  GI:3344677
KEYWORDS    EST.
SOURCE      fission yeast.
  ORGANISM  Schizosaccharomyces pombe
            Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
            Schizosaccharomycetales; Schizosaccharomycetaceae;
            Schizosaccharomyces.
REFERENCE   1  (bases 1 to 62)
  AUTHORS   Morimyo,M. and Mita,K.
  TITLE     Identification of expressed sequence tags of Schizosaccharomyces
            pombe
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mitsuoki Morimyo
            Genome Research Group
            National Institute of Radiological Sciences
            9-1, Anagawa-4-chome, Inage-ku, Chiba, Chiba 263-8555, Japan
            Email: morimyo@nirs.go.jp.
FEATURES             Location/Qualifiers
     source          1. .62
                     /organism="Schizosaccharomyces pombe"
                     /strain="972"
                     /db_xref="taxon:4896"
                     /clone="spc03066"
                     /clone_lib="Schizosaccharomyces pombe late log phase cDNA"
                     /sex="h minus"
                     /note="Vector: M13mp19; The cDNA library of
                     Schizosaccharomyces pombe was prepared by cloning cDNA
                     into the SmaI site of M13mp19 DNA and the direction of DNA
                     sequences was not always from 5' to 3'. The cDNA data of
                     Schizosaccharomyces pombe are available for searching on
                     the World Wide Web.  (URL, http://www.nirs.go.jp)"
BASE COUNT       18 a      13 c      11 g      20 t
ORIGIN
```

```
  Query Match             100.0%;  Score 8;  DB 10;  Length 62;
  Best Local Similarity   100.0%;  Pred. No. 3.2e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 AACGTTCG 8
          ||||||||
Db     32 AACGTTCG 39


RESULT   7
AU008222
LOCUS       AU008222       62 bp    mRNA            EST       31-JUL-1998
DEFINITION  AU008222 Schizosaccharomyces pombe late log phase cDNA
            Schizosaccharomyces pombe cDNA clone spc03071, mRNA sequence.
ACCESSION   AU008222
VERSION     AU008222.1  GI:3344680
KEYWORDS    EST.
SOURCE      fission yeast.
  ORGANISM  Schizosaccharomyces pombe
            Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
            Schizosaccharomycetales; Schizosaccharomycetaceae;
            Schizosaccharomyces.
REFERENCE   1  (bases 1 to 62)
  AUTHORS   Morimyo,M. and Mita,K.
  TITLE     Identification of expressed sequence tags of Schizosaccharomyces
            pombe
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mitsuoki Morimyo
            Genome Research Group
            National Institute of Radiological Sciences
            9-1, Anagawa-4-chome, Inage-ku, Chiba, Chiba 263-8555, Japan
            Email: morimyo@nirs.go.jp.
FEATURES             Location/Qualifiers
     source          1. .62
                     /organism="Schizosaccharomyces pombe"
                     /strain="972"
                     /db_xref="taxon:4896"
                     /clone="spc03071"
                     /clone_lib="Schizosaccharomyces pombe late log phase cDNA"
                     /sex="h minus"
                     /note="Vector: M13mp19; The cDNA library of
                     Schizosaccharomyces pombe was prepared by cloning cDNA
                     into the SmaI site of M13mp19 DNA and the direction of DNA
                     sequences was not always from 5' to 3'. The cDNA data of
                     Schizosaccharomyces pombe are available for searching on
                     the World Wide Web.  (URL, http://www.nirs.go.jp)"
BASE COUNT       18 a     13 c     11 g     20 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 62;
  Best Local Similarity   100.0%;  Pred. No. 3.2e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 AACGTTCG 8
          ||||||||
Db     32 AACGTTCG 39


RESULT   8
AU008233
LOCUS       AU008233       62 bp    mRNA            EST       31-JUL-1998
DEFINITION  AU008233 Schizosaccharomyces pombe late log phase cDNA
            Schizosaccharomyces pombe cDNA clone spc03087, mRNA sequence.
ACCESSION   AU008233
VERSION     AU008233.1  GI:3344691
KEYWORDS    EST.
SOURCE      fission yeast.
  ORGANISM  Schizosaccharomyces pombe
            Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
            Schizosaccharomycetales; Schizosaccharomycetaceae;
            Schizosaccharomyces.
REFERENCE   1  (bases 1 to 62)
  AUTHORS   Morimyo,M. and Mita,K.
  TITLE     Identification of expressed sequence tags of Schizosaccharomyces
            pombe
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mitsuoki Morimyo
            Genome Research Group
            National Institute of Radiological Sciences
            9-1, Anagawa-4-chome, Inage-ku, Chiba, Chiba 263-8555, Japan
            Email: morimyo@nirs.go.jp.
FEATURES             Location/Qualifiers
     source          1. .62
                     /organism="Schizosaccharomyces pombe"
                     /strain="972"
                     /db_xref="taxon:4896"
                     /clone="spc03087"
                     /clone_lib="Schizosaccharomyces pombe late log phase cDNA"
                     /sex="h minus"
                     /note="Vector: M13mp19; The cDNA library of
                     Schizosaccharomyces pombe was prepared by cloning cDNA
                     into the SmaI site of M13mp19 DNA and the direction of DNA
                     sequences was not always from 5' to 3'. The cDNA data of
                     Schizosaccharomyces pombe are available for searching on
                     the World Wide Web.  (URL, http://www.nirs.go.jp)"
BASE COUNT       18 a     13 c     11 g     20 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 62;
  Best Local Similarity   100.0%;  Pred. No. 3.2e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 AACGTTCG 8
          ||||||||
Db     32 AACGTTCG 39


RESULT   9
AU008237
LOCUS       AU008237       62 bp    mRNA            EST       31-JUL-1998
DEFINITION  AU008237 Schizosaccharomyces pombe late log phase cDNA
            Schizosaccharomyces pombe cDNA clone spc03091, mRNA sequence.
ACCESSION   AU008237
VERSION     AU008237.1  GI:3344695
KEYWORDS    EST.
SOURCE      fission yeast.
  ORGANISM  Schizosaccharomyces pombe
            Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
            Schizosaccharomycetales; Schizosaccharomycetaceae;
            Schizosaccharomyces.
REFERENCE   1  (bases 1 to 62)
  AUTHORS   Morimyo,M. and Mita,K.
  TITLE     Identification of expressed sequence tags of Schizosaccharomyces
            pombe
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mitsuoki Morimyo
            Genome Research Group
            National Institute of Radiological Sciences
            9-1, Anagawa-4-chome, Inage-ku, Chiba, Chiba 263-8555, Japan
            Email: morimyo@nirs.go.jp.
FEATURES             Location/Qualifiers
     source          1. .62
                     /organism="Schizosaccharomyces pombe"
                     /strain="972"
                     /db_xref="taxon:4896"
                     /clone="spc03091"
                     /clone_lib="Schizosaccharomyces pombe late log phase cDNA"
                     /sex="h minus"
                     /note="Vector: M13mp19; The cDNA library of
                     Schizosaccharomyces pombe was prepared by cloning cDNA
                     into the SmaI site of M13mp19 DNA and the direction of DNA
                     sequences was not always from 5' to 3'. The cDNA data of
                     Schizosaccharomyces pombe are available for searching on
```

```
                    the World Wide Web.  (URL, http://www.nirs.go.jp)"
BASE COUNT       18 a     13 c     11 g     20 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 62;
  Best Local Similarity   100.0%;  Pred. No. 3.2e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      32 AACGTTCG 39


RESULT  10
BE638333
LOCUS       BE638333       64 bp    mRNA            EST       28-AUG-2000
DEFINITION  SWOvMfCAR18A08SK Onchocerca volvulus microfilaria cDNA
            (SAW98MLW-OvMf) Onchocerca volvulus cDNA clone SWOvMfCAR18A08 5',
            mRNA sequence.
ACCESSION   BE638333
VERSION     BE638333.1  GI:9937035
KEYWORDS    EST.
SOURCE      Onchocerca volvulus.
  ORGANISM  Onchocerca volvulus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
            Onchocercidae; Onchocerca.
REFERENCE   1  (bases 1 to 64)
  AUTHORS   Williams,S.A.
  TITLE     Genes expressed in microfilaria of Onchocerca volvulus
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Steven A. Williams
            Molecular Parasitology
            Smith College Department of Biological Sciences
            Department of Biological Sciences, Clark Science Center, Smith
            College, Northampton, MA, 01063, USA
            Tel: 4135853826
            Fax: 4135853786
            Email: genome@smith.edu
            Seq primer: pBluescript SK.
FEATURES             Location/Qualifiers
     source          1. .64
                     /organism="Onchocerca volvulus"
                     /db_xref="taxon:6282"
                     /clone="SWOvMfCAR18A08"
                     /clone_lib="Onchocerca volvulus microfilaria cDNA
                     (SAW98MLW-OvMf)"
                     /dev_stage="microfilaria"
                     /lab_host="XL1-Blue MRF'"
                     /note="Vector: Lambda Uni-ZAP XR; Site_1: Eco RI; Site_2:
                     Xho I; Filarial nematode parasite of humans. mRNA was
                     prepared from approximately 200,000 microfilariae isolated
                     from  the  skin of infected individuals from Kumba,
                     Cameroon and converted to  double-stranded cDNA using
                     reverse transcriptase and oligo(dT) followed by RNase H
                     and DNA pol I. The library has 7.8 x 10E4 independent
                     recombinants and the average insert size is  approximately
                     1kb. The library was constructed by Michelle
                     Lizotte-Waniewski. The library is available from
                     Dr.S.A.Williams,  email:genome@smith.edu."
BASE COUNT       22 a     11 c     19 g     12 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 64;
  Best Local Similarity   100.0%;  Pred. No. 3.2e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       5 AACGTTCG 12


RESULT  11
BI097404
LOCUS       BI097404       64 bp    mRNA            EST       25-JUN-2001
DEFINITION  SWOv3MCAM63D09SK Onchocerca volvulus molting L3 larva cDNA
            (SL96MLW-OvmL3) Onchocerca volvulus cDNA clone SWOv3MCAM63D09 5',
            mRNA sequence.
ACCESSION   BI097404
VERSION     BI097404.1  GI:14549061
KEYWORDS    EST.
SOURCE      Onchocerca volvulus.
  ORGANISM  Onchocerca volvulus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
            Onchocercidae; Onchocerca.
REFERENCE   1  (bases 1 to 64)
  AUTHORS   Williams,S.A., Lizotte-Waniewski,M., Laney,S. and Lustigman,S.
  TITLE     Genes expressed in molting L3 larvae of Onchocerca volvulus
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Steven A. Williams
            Molecular Parasitology
            Smith College Department of Biological Sciences
            Department of Biological Sciences, Clark Science Center, Smith
            College, Northampton, MA, 01063, USA
            Tel: 4135853826
            Fax: 4135853786
            Email: genome@smith.edu
            Seq primer: pBluescript SK.
FEATURES             Location/Qualifiers
     source          1. .64
                     /organism="Onchocerca volvulus"
                     /strain="Kumba, Cameroons"
                     /db_xref="taxon:6282"
                     /clone="SWOv3MCAM63D09"
                     /clone_lib="Onchocerca volvulus molting L3 larva cDNA
                     (SL96MLW-OvmL3)"
                     /dev_stage="molting L3"
                     /lab_host="XL1-Blue MRF'"
                     /note="Vector: Lambda Uni-ZAP XR; Site_1: Eco RI; Site_2:
                     Xho I; Filarial nematode parasite of humans. Third-stage
                     larvae, L3,  were isolated from infected black flies in
                     Cameroon (forest strain). The L3 were cultured in 20% FCS
                     in IMDM+ NCTC 135 and collected after day 1, 2, or 3 in
                     culture. L3 of O. volvulus molt to fourth-stage larvae by
                     day 5 in culture. mRNA was isolated from approximately
                     6000 molting larvae (mL3),2000 larvae from day 1, 2 or 3
                     in culture, and converted to double-stranded cDNA using
                     reverse transcriptase and oligo(dT) followed by RNase H
                     and DNA pol I. The library was constructed in the lambda
                     Uni-Zap XR vector and has 1 x 10E6 independent
                     recombinants and the average insert size is ~1200 bp. The
                     library was constructed by Sara Lustigman and Michelle
                     Lizotte-Waniewski in the Laboratory of Dr. S. A. Williams.
                     The library is available from Dr. Sara Lustigman (email:
                     slustigm@nybc.org)."
BASE COUNT       17 a     11 c     19 g     17 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 11;  Length 64;
  Best Local Similarity   100.0%;  Pred. No. 3.2e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      16 AACGTTCG 23


RESULT  12
AA617006/c
LOCUS       AA617006       67 bp    mRNA            EST       07-OCT-1997
DEFINITION  vk51a11.r1 Stratagene mouse Tcell 937311 Mus musculus cDNA clone
            IMAGE:958172 5' similar to gb:U12403 Mus musculus Csa-19 mRNA,
            complete cds (MOUSE);, mRNA sequence.
ACCESSION   AA617006
```

```
VERSION     AA617006.1  GI:2504211
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 67)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:546964
            Seq primer: -28ml3 rev1 ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .67
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:958172"
                     /clone_lib="Stratagene mouse Tcell 937311"
                     /tissue_type="Tcell"
                     /dev_stage="M30 CD4+ cells"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Organ: blood; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                     Oligo dT. M30 CD4+ cells. Average insert size: 1.0 kb;
                     Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCACGAG
                     3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT      23 a      14 c      15 g      15 t
ORIGIN
```

```
  Query Match           100.0%;  Score 8;  DB 10;  Length 67;
  Best Local Similarity  100.0%;  Pred. No. 3.3e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
          ||||||||
Db      66 AACGTTCG 59
```

```
RESULT  13
AQ025258/c
LOCUS       AQ025258        69 bp    DNA             GSS       23-AUG-2000
DEFINITION  EP(3)3076 Drosophila melanogaster EP line Drosophila melanogaster
            genomic Sequence recovered from 5' end of P element, DNA sequence.
ACCESSION   AQ025258
VERSION     AQ025258.1  GI:3265610
KEYWORDS    GSS.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 69)
  AUTHORS   Liao,G.-C., Rehm,E.J. and Rubin,G.M.
  TITLE     Insertion site preferences of the P transposable element in
            Drosophila melanogaster
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3347-3351 (2000)
  MEDLINE   20202638
COMMENT     Contact: Gerald Rubin
            Berkeley Drosophila Genome Project
            University of California, Berkeley
            LSA Building, Berkeley, CA 94720-3200, USA
            Fax: 5106439947
            Email: gerry@fruitfly.berkeley.edu
            Sequence recovery method was inverse PCR.

            Sequence orientation is forward strand relative to 5' end of P
            element

            The P element insertion position is base 62 in the 69 bases. This
            insertion position refers to the first base of the 8 base target
            recognition sequence.
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .69
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone_lib="Drosophila melanogaster EP line"
                     /note="Inverse PCR was performed on Drosophila
                     melanogaster strains each of which contains a single EP
                     transposable element insertion. (The generation of these
                     insertion strains is described in Rorth P, Szabo K, Bailey
                     A, Laverty T, Rehm J, Rubin GM, Weigmann K, Milan M, Benes
                     V, Ansorge W, Cohen SM. 1998. Systematic gain-of-function
                     genetics in Drosophila. Development 6:1049-1057.) The
                     resultant fragment for each strain was directly sequenced
                     to determine the genomic sequence at the site of
                     insertion. Details of the protocols used can be found at
                     http://fruitfly.berkeley.edu/p_disrupt/inverse_pcr.html."
BASE COUNT      23 a      12 c      15 g      17 t       2 others
ORIGIN
```

```
  Query Match           100.0%;  Score 8;  DB 13;  Length 69;
  Best Local Similarity  100.0%;  Pred. No. 3.3e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
          ||||||||
Db      40 AACGTTCG 33
```

```
RESULT  14
BE024070/c
LOCUS       BE024070        71 bp    mRNA            EST       31-JUL-2001
DEFINITION  sm96c11.y1 Gm-c1015 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
            Gm-c1015-7917 5', mRNA sequence.
ACCESSION   BE024070
VERSION     BE024070.1  GI:8286511
KEYWORDS    EST.
SOURCE      soybean.
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 71)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
            ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
            Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
            ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
            ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
            ,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
```

This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomesystems.com web site: www.genomesystems.com
Insert Length: 264 Std Error: 0.00
Seq primer: -40RP from Gibco.

```
FEATURES             Location/Qualifiers
     source          1. .71
                     /organism="Glycine max"
                     /db_xref="taxon:3847"
                     /clone="GENOME SYSTEMS CLONE ID: Gm-c1015-7917"
                     /clone_lib="Gm-c1015"
                     /tissue_type="Mature flowers, field grown plants"
                     /lab_host="XL10-Gold"
                     /note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2:
                     XhoI; This cDNA library was constructed from mRNA isolated
                     from mature flowers of field grown plants. The cDNA
                     library was prepared using the Stratagene pBluescript II
                     XR cDNA library construction kit. Complementary DNA was
                     synthesized from mRNA using a primer consisting of a poly
                     (dT) sequence with a XhoI restriction site. EcoRI adapters
                     were ligated to the blunt-ended cDNA fragments followed by
                     XhoI digestion. The cDNA fragments were directionally
                     cloned into the EcoRI-XhoI restriction site of the
                     pBluescript vector. The ligated cDNA fragments were
                     transformed into XL10-Gold host cells. This library was
                     constructed by Dr. Randy Shoemaker and Dr. John
                     Erpelding."
BASE COUNT       17 a      8 c     14 g     32 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 71;
  Best Local Similarity   100.0%;  Pred. No. 3.3e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      54 AACGTTCG 47


RESULT  15
AW499129
LOCUS       AW499129       73 bp    mRNA            EST       01-MAR-2000
DEFINITION  SWOvAFCAP39G07SK Onchocerca volvulus adult female cDNA
            (SAW98MLW-OvAF) Onchocerca volvulus cDNA clone SWOvAFCAP39G07 5',
            mRNA sequence.
ACCESSION   AW499129
VERSION     AW499129.1  GI:7137509
KEYWORDS    EST.
SOURCE      Onchocerca volvulus.
  ORGANISM  Onchocerca volvulus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
            Onchocercidae; Onchocerca.
REFERENCE   1  (bases 1 to 73)
  AUTHORS   Lizotte-Waniewski,M. and Williams,S.A.
  TITLE     Genes expressed in adult female stage of Onchocerca volvulus
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Steven A. Williams
            Molecular Parasitology
            Smith College Department of Biological Sciences
            Department of Biological Sciences, Clark Science Center, Smith
            College, Northampton, MA, 01063, USA
            Tel: 4135853826
            Fax: 4135853786
            Email: genome@smith.edu
            Seq primer: pBluescript SK.
FEATURES             Location/Qualifiers
     source          1. .73
                     /organism="Onchocerca volvulus"
                     /db_xref="taxon:6282"
                     /clone="SWOvAFCAP39G07"
                     /clone_lib="Onchocerca volvulus adult female cDNA
                     (SAW98MLW-OvAF)"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="XL1-Blue MRF'"
                     /note="Vector: Lambda Uni-ZAP XR; Site_1: Eco RI; Site_2:
                     Xho I; Filarial nematode parasite of humans. Two adult
                     female worms of Onchocerca volvulus were isolated from
                     consenting patients and quick frozen. Adult female mRNA
                     was converted to double-stranded cDNA using reverse
                     transcriptase and oligo(dT) followed by RNase H and DNA
                     pol I. The library has 7 x 10E5 independent recombinants
                     and the average insert size is ~1100bp.The library was
                     constructed by Michelle Lizotte-Waniewski with worms
                     provided by Dr. Sara Lustigman. The library is available
                     from Dr.Steven A. Williams, email: genome@smith.edu."
BASE COUNT       21 a     14 c     22 g     16 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 73;
  Best Local Similarity   100.0%;  Pred. No. 3.3e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      66 AACGTTCG 73


RESULT  16
AA404533
LOCUS       AA404533       74 bp    mRNA            EST       17-MAY-1997
DEFINITION  zw37h02.s1 Soares_total_fetus_Nb2HF8_9w Homo sapiens cDNA clone
            IMAGE:772275 3' similar to gb:A18658 INSULIN RECEPTOR PRECURSOR
            (HUMAN);, mRNA sequence.
ACCESSION   AA404533
VERSION     AA404533.1  GI:2059283
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 74)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.
            , Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie
            ,T., Waterston,R. and Wilson,R.
  TITLE     WashU-Merck EST Project 1997
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Seq primer: -41m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .74
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:772275"
                     /clone_lib="Soares_total_fetus_Nb2HF8_9w"
                     /dev_stage="8-9 weeks"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from mRNA obtained from pooled 8-9 week
                     (total) fetus material with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCTTAATTTTTTTTTTTTTTTT 3']
```

Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 12 a 9 c 20 g 33 t
ORIGIN

Query Match 100.0%; Score 8; DB 10; Length 74;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACGTTCG 8
||||||||
Db 58 AACGTTCG 65

RESULT 17
BE027432
LOCUS BE027432 76 bp mRNA EST 07-JUN-2000
DEFINITION EtESTea43g05.y1 Eimeria M5-6 Merozoite stage subtracted Eimeria tenella cDNA 5' similar to SW:TA4_EIMTE P13399 SPORULATED OOCYST TA4 ANTIGEN PRECURSOR ;, mRNA sequence.
ACCESSION BE027432
VERSION BE027432.1 GI:8320802
KEYWORDS EST.
SOURCE Eimeria tenella.
ORGANISM Eimeria tenella
Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Eimeriidae; Eimeria.
REFERENCE 1 (bases 1 to 76)
AUTHORS Liberator,P., Diaz,C., Tang,K., Marra,M., Hillier,L., Kucaba,T., Martin,J., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen ,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey ,N., Schurk,R., Ritter,E., Kohn,S., Florence,N., Shin,T., Jackson ,Y., Cardenas,M., McCann,R., Waterston,R., Wilson,R. and Sibley,D.
TITLE WashU-Merck Eimeria tenella project
JOURNAL Unpublished (1999)
COMMENT Contact: David Sibley, Ph.D.
WashU-Merck Eimeria tenella project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Contact David Sibley (toxoest@borcim.wustl.edu) for further information relating to organism, libraries, or clone availability.
Seq primer: -40RP from Gibco.
FEATURES Location/Qualifiers
source 1. .76
/organism="Eimeria tenella"
/strain="LS18"
/db_xref="taxon:5802"
/clone_lib="Eimeria M5-6 Merozoite stage subtracted"
/dev_stage="Merozoite"
/lab_host="SOLR E. coli"
/note="Vector: Bluescript SK-; Site_1: EcoRI; Site_2: XhoI ; Merozoites were obtained from ceacal scrapings of chickens infected with E. tenella. cDNA was synthesized from poly mRNA using an oligo-dT primer containing a XhoI site. Following second strand synthesis, EcoRI adapters were ligated to the cDNA and products were size-selected on Sephacryl S500. The cDNAs were ligated to EcoRI/XhoI prepared lambda ZAPII(Stratagene). Clones were converted to phagemids by mass excision using ExAssist helper phage and E.coli SOLR cells (Stratagene). Insert sizes range from 0.7-1.5Kb. The library may contain a small percentage of host or bacterial contaminants. Clones were selected by negative hybridization against a pool of over-represented ESTs (N>=10, from 1506 previous reads)."
BASE COUNT 27 a 14 c 22 g 13 t
ORIGIN

Query Match 100.0%; Score 8; DB 10; Length 76;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACGTTCG 8
||||||||
Db 67 AACGTTCG 74

RESULT 18
TA389C08P/c
LOCUS TA389C08P 80 bp DNA GSS 13-DEC-2000
DEFINITION T. brucei sheared genomic DNA clone 389c08, forward sequence, genomic survey sequence.
ACCESSION AL498967
VERSION AL498967.1 GI:11874689
KEYWORDS GSS.
SOURCE Trypanosoma brucei.
ORGANISM Trypanosoma brucei
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
REFERENCE 1 (bases 1 to 80)
AUTHORS Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R., Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L., Melville,S.E., Rajandream,M.A. and Barrell,B.G.
TITLE Direct Submission
JOURNAL Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and nh1@sanger.ac.uk
COMMENT Constructed at the Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from a cloned population of Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared to give a tight size distribution (
4 kb). The v + i method used for the library construction is described in detail in Smith, H. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999).
Email: nelsayed@tigr.org
Details of T. brucei sequencing at the Sanger Centre are available at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES Location/Qualifiers
source 1. .80
/organism="Trypanosoma brucei"
/strain="TREU927"
/db_xref="taxon:5691"
/clone="389c08"
BASE COUNT 14 a 25 c 17 g 24 t
ORIGIN

Query Match 100.0%; Score 8; DB 13; Length 80;
Best Local Similarity 100.0%; Pred. No. 3.4e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACGTTCG 8
||||||||
Db 8 AACGTTCG 1

RESULT 19
AI903642
LOCUS AI903642 81 bp mRNA EST 30-MAR-2000
DEFINITION QV-BT032-190299-085_2 BT032 Homo sapiens cDNA, mRNA sequence.
ACCESSION AI903642
VERSION AI903642.1 GI:6494029
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 81)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/seq/gethtml.pl?t1=QV&t2=QV-BT032-085_2.ht
            ml&t3=190299&t4=1)
            Seq primer: puc 18 forward.
FEATURES              Location/Qualifiers
     source           1. .81
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone_lib="BT032"
                      /sex="female"
                      /dev_stage="Adult"
                      /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
                      SmaI; A mini-library was made by cloning products derived
                      from  ORESTES PCR (U.S. Letters Patent application No. 196
                      ,716 -  Ludwig Institute for Cancer Research) profiles
                      into  the pUC 18 vector. Reverse transcription of tissue
                      mRNA and cDNA amplification were performed under low
                      stringency conditions."
BASE COUNT       23 a     25 c     25 g      8 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 81;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      65 AACGTTCG 72


RESULT  20
BE027387
LOCUS       BE027387       81 bp    mRNA            EST       07-JUN-2000
DEFINITION  EtESTea43b04.y1 Eimeria M5-6 Merozoite stage subtracted Eimeria
            tenella cDNA 5' similar to SW:TA4_EIMTE P13399 SPORULATED OOCYST
            TA4 ANTIGEN PRECURSOR ;, mRNA sequence.
ACCESSION   BE027387
VERSION     BE027387.1  GI:8320753
KEYWORDS    EST.
SOURCE      Eimeria tenella.
  ORGANISM  Eimeria tenella
            Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Eimeriidae;
            Eimeria.
REFERENCE   1  (bases 1 to 81)
  AUTHORS   Liberator,P., Diaz,C., Tang,K., Marra,M., Hillier,L., Kucaba,T.,
            Martin,J., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen
            ,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey
            ,N., Schurk,R., Ritter,E., Kohn,S., Florence,N., Shin,T., Jackson
            ,Y., Cardenas,M., McCann,R., Waterston,R., Wilson,R. and Sibley,D.
  TITLE     WashU-Merck Eimeria tenella project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: David Sibley, Ph.D.
            WashU-Merck Eimeria tenella project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Contact David Sibley (toxoest@borcim.wustl.edu) for further
            information relating to organism, libraries, or clone availability.
            Seq primer: -40RP from Gibco.
FEATURES              Location/Qualifiers
     source           1. .81
                      /organism="Eimeria tenella"
                      /strain="LS18"
                      /db_xref="taxon:5802"
                      /clone_lib="Eimeria M5-6 Merozoite stage subtracted"
                      /dev_stage="Merozoite"
                      /lab_host="SOLR E. coli"
                      /note="Vector: Bluescript SK-; Site_1: EcoRI; Site_2: XhoI
                      ; Merozoites were obtained from ceacal scrapings of
                      chickens infected with E. tenella. cDNA was synthesized
                      from poly mRNA using an oligo-dT primer containing a XhoI
                      site. Following second strand synthesis, EcoRI adapters
                      were ligated to the cDNA and products were size-selected
                      on Sephacryl S500. The cDNAs were ligated to EcoRI/XhoI
                      prepared lambda ZAPII(Stratagene). Clones were converted
                      to phagemids by mass excision using ExAssist helper phage
                      and E.coli SOLR cells (Stratagene). Insert sizes range
                      from 0.7-1.5Kb. The library may contain a small percentage
                      of host or bacterial contaminants. Clones were selected by
                      negative hybridization against a pool of  over-represented
                      ESTs (N>=10, from 1506 previous reads)."
BASE COUNT       30 a     15 c     22 g     14 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 81;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      49 AACGTTCG 56


RESULT  21
AA629864
LOCUS       AA629864       85 bp    mRNA            EST       06-MAR-1998
DEFINITION  ad48h11.s1 Stratagene lung carcinoma 937218 Homo sapiens cDNA clone
            IMAGE:884997 3' similar to TR:E196749 E196749 MRNA; EXPRESSED
            SEQUENCE TAG ;, mRNA sequence.
ACCESSION   AA629864
VERSION     AA629864.1  GI:2552475
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 85)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
```

```
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Insert Length: 899    Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES                Location/Qualifiers
     source             1. .85
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone="IMAGE:884997"
                        /clone_lib="Stratagene lung carcinoma 937218"
                        /tissue_type="lung carcinoma"
                        /cell_line="NCI-H69"
                        /dev_stage="cell line NCI-H69"
                        /lab_host="SOLR (kanamycin resistant)"
                        /note="Organ: lung; Vector: pBluescript SK-; Site_1: EcoRI
                        ; Site_2: XhoI; Cloned unidirectionally.  Primer: Oligo
                        dT. Small cell carcinoma cell line NCI-H69.  Average
                        insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor
                        sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor sequence: 5'
                        CTCGAGTTTTTTTTTTTTTTTTT 3'"
BASE COUNT       19 a      24 c      22 g      20 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 85;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      68 AACGTTCG 75


RESULT  22
AA629864/c
LOCUS       AA629864       85 bp    mRNA            EST       06-MAR-1998
DEFINITION  ad48h11.s1 Stratagene lung carcinoma 937218 Homo sapiens cDNA clone
            IMAGE:884997 3' similar to TR:E196749 E196749 MRNA; EXPRESSED
            SEQUENCE TAG ;, mRNA sequence.
ACCESSION   AA629864
VERSION     AA629864.1  GI:2552475
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 85)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Insert Length: 899    Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES                Location/Qualifiers
     source             1. .85
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone="IMAGE:884997"
                        /clone_lib="Stratagene lung carcinoma 937218"
                        /tissue_type="lung carcinoma"
                        /cell_line="NCI-H69"
                        /dev_stage="cell line NCI-H69"
                        /lab_host="SOLR (kanamycin resistant)"
                        /note="Organ: lung; Vector: pBluescript SK-; Site_1: EcoRI
                        ; Site_2: XhoI; Cloned unidirectionally.  Primer: Oligo
                        dT. Small cell carcinoma cell line NCI-H69.  Average
                        insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor
                        sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor sequence: 5'
                        CTCGAGTTTTTTTTTTTTTTTTT 3'"
BASE COUNT       19 a      24 c      22 g      20 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 85;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      73 AACGTTCG 66


RESULT  23
AA670169
LOCUS       AA670169       85 bp    mRNA            EST       20-NOV-1997
DEFINITION  ab65d05.s1 Stratagene lung carcinoma 937218 Homo sapiens cDNA clone
            IMAGE:845673 3' similar to TR:E196749 E196749 MRNA; EXPRESSED
            SEQUENCE TAG ;, mRNA sequence.
ACCESSION   AA670169
VERSION     AA670169.1  GI:2631668
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 85)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES                Location/Qualifiers
     source             1. .85
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone="IMAGE:845673"
                        /clone_lib="Stratagene lung carcinoma 937218"
                        /tissue_type="lung carcinoma"
                        /cell_line="NCI-H69"
                        /dev_stage="cell line NCI-H69"
                        /lab_host="SOLR (kanamycin resistant)"
                        /note="Organ: lung; Vector: pBluescript SK-; Site_1: EcoRI
                        ; Site_2: XhoI; Cloned unidirectionally.  Primer: Oligo
                        dT. Small cell carcinoma cell line NCI-H69.  Average
                        insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor
                        sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor sequence: 5'
                        CTCGAGTTTTTTTTTTTTTTTTT 3'"
BASE COUNT       19 a      26 c      24 g      16 t
ORIGIN
```

```
 Query Match             100.0%;  Score 8;  DB 10;  Length 85;
 Best Local Similarity   100.0%;  Pred. No. 3.4e+04;
 Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      68 AACGTTCG 75


RESULT  24
AA670169/c
LOCUS       AA670169       85 bp    mRNA            EST       20-NOV-1997
DEFINITION  ab65d05.s1 Stratagene lung carcinoma 937218 Homo sapiens cDNA clone
            IMAGE:845673 3' similar to TR:E196749 E196749 MRNA; EXPRESSED
            SEQUENCE TAG ;, mRNA sequence.
ACCESSION   AA670169
VERSION     AA670169.1  GI:2631668
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 85)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .85
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:845673"
                     /clone_lib="Stratagene lung carcinoma 937218"
                     /tissue_type="lung carcinoma"
                     /cell_line="NCI-H69"
                     /dev_stage="cell line NCI-H69"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Organ: lung; Vector: pBluescript SK-; Site_1: EcoRI
                     ; Site_2: XhoI; Cloned unidirectionally.  Primer: Oligo
                     dT. Small cell carcinoma cell line NCI-H69.  Average
                     insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor
                     sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor sequence: 5'
                     CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT       19 a     26 c     24 g     16 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 85;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      73 AACGTTCG 66


RESULT  25
TA245G05Q/c
LOCUS       TA245G05Q      86 bp    DNA             GSS       13-DEC-2000
DEFINITION  T. brucei sheared genomic DNA clone 245g05, reverse sequence,
            genomic survey sequence.
ACCESSION   AL482195
VERSION     AL482195.1  GI:11848200
KEYWORDS    GSS.
SOURCE      Trypanosoma brucei.
  ORGANISM  Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 86)
  AUTHORS   Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
            Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
            Melville,S.E., Rajandream,M.A. and Barrell,B.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
            project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
            Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
            nh1@sanger.ac.uk
COMMENT     Constructed at the Institute for Genomic Research (TIGR),
            Rockville, MD. Genomic DNA isolated from a cloned population of
            Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
            to give a tight size distribution (
            4 kb).  The v + i method used for the library construction is
            described in detail in Smith, H. and Venter, J.C. (Making small
            insert libraries for whole genome shotgun sequencing projects.. In
            Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
            Barrell, Oxford University Press, 1999).
            Email: nelsayed@tigr.org
            Details of T. brucei sequencing at the Sanger Centre are available
            at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES             Location/Qualifiers
     source          1. .86
                     /organism="Trypanosoma brucei"
                     /strain="TREU927"
                     /db_xref="taxon:5691"
                     /clone="245g05"
BASE COUNT       38 a     12 c     17 g     19 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 13;  Length 86;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      12 AACGTTCG 5


RESULT  26
AI289175/c
LOCUS       AI289175       88 bp    mRNA            EST       01-FEB-1999
DEFINITION  qn25f09.x1 NCI_CGAP_Kid5 Homo sapiens cDNA clone IMAGE:1899305 3',
            mRNA sequence.
ACCESSION   AI289175
VERSION     AI289175.1  GI:3932439
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 88)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
```

```
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 2517    Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 69.
FEATURES                Location/Qualifiers
     source             1. .88
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone="IMAGE:1899305"
                        /clone_lib="NCI_CGAP_Kid5"
                        /tissue_type="2 pooled tumors (clear cell type)"
                        /lab_host="DH10B"
                        /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                        a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                        strand cDNA was primed with a Not I - oligo(dT) primer [5'
                        AACTGGAAGAATTCGCGGCCGCAATATTTTTTTTTTTTTTTT 3'],
                        double-stranded cDNA was ligated to Eco RI adaptors
                        (Pharmacia), digested with Not I and cloned into the Not I
                        and Eco RI sites of the modified pT7T3 vector. Library
                        went through one round of normalization. Library
                        constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT        15 a      24 c      12 g      37 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 88;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      16 AACGTTCG 9


RESULT  27
AA626216/c
LOCUS       AA626216       88 bp    mRNA            EST       15-OCT-1997
DEFINITION  zv88a05.s1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:766832 3'
            similar to TR:G300372 G300372 CELL GROWTH REGULATING NUCLEOLAR
            PROTEIN. ;, mRNA sequence.
ACCESSION   AA626216
VERSION     AA626216.1  GI:2538603
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 88)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -40m13 fwd. ET from Amersham
FEATURES    High quality sequence stop: 1.
     source         Location/Qualifiers
                    1. .88
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:766832"
                    /clone_lib="Soares_NhHMPu_S1"
                    /tissue_type="Pooled human melanocyte, fetal heart, and
                    pregnant uterus"
                    /lab_host="DH10B"
                    /note="Organ: mixed (see below); Vector: pT7T3D-Pac
                    (Pharmacia) with a modified polylinker; Site_1: Not I;
                    Site_2: Eco RI; Equal amounts of plasmid DNA from three
                    normalized libraries (melanocyte 2NbHM, pregnant uterus
                    NbHPU, and fetal heart NbHH19W) were mixed, and ss circles
                    were made in vitro. Following HAP purification, this DNA
                    was used as tracer in a subtractive hybridization
                    reaction. The driver was PCR-amplified cDNAs from pools of
                    5,000 clones made from the same 3 libraries. The pools
                    consisted of I.M.A.G.E. clones 260232-265223,
                    340488-345479, and 484488-489479."
BASE COUNT        23 a      22 c      19 g      24 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 88;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      34 AACGTTCG 27


RESULT  28
AJ239919/c
LOCUS       AJ239919       94 bp    mRNA            EST       10-AUG-1999
DEFINITION  AJ239919 Aspergillus niger ATCC6275 Aspergillus niger cDNA clone
            AN06D12, mRNA sequence.
ACCESSION   AJ239919
VERSION     AJ239919.1  GI:5443910
KEYWORDS    EST.
SOURCE      Aspergillus niger.
  ORGANISM  Aspergillus niger
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
            Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
REFERENCE   1  (bases 1 to 94)
  AUTHORS   Choi,J.Y., Lee,D.W., Koh,J.S., Kim,J.H., Yang,M.S. and Chae,K.S.
  TITLE     Identification of expressed sequence tags (ESTs) of the highly
            transcribed genes in Aspergillus niger
  JOURNAL   Biotechnol. Lett. 21, 381-384 (1999)
COMMENT     Contact: Chae KS
            Faculty of Biological Sciences
            Chonbuk National University
            Chonju 561-756, Republic of Korea.
FEATURES                Location/Qualifiers
     source             1. .94
                        /organism="Aspergillus niger"
                        /strain="ATCC6275"
                        /db_xref="taxon:5061"
                        /clone="AN06D12"
                        /clone_lib="Aspergillus niger ATCC6275"
BASE COUNT        30 a      19 c      17 g      22 t      6 others
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 94;
  Best Local Similarity   100.0%;  Pred. No. 3.5e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      52 AACGTTCG 45


RESULT  29
BE576515
```

```
LOCUS       BE576515       94 bp    mRNA            EST       15-AUG-2000
DEFINITION  dc40g03.y1 NICHD XGC Emb3 Xenopus laevis cDNA clone IMAGE:3399604
            5', mRNA sequence.
ACCESSION   BE576515
VERSION     BE576515.1  GI:9826314
KEYWORDS    EST.
SOURCE      African clawed frog.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 94)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Other_ESTs: dc40g03.x1
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Martha Rebbert, Steven L. Klein, Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: Xenopus clones from this library are available
            through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
            Seq primer: -40RP from Gibco
            High quality sequence stop: 83.
FEATURES             Location/Qualifiers
     source          1. .94
                     /organism="Xenopus laevis"
                     /db_xref="taxon:8355"
                     /clone="IMAGE:3399604"
                     /clone_lib="NICHD XGC Emb3"
                     /tissue_type="embryo (stages 24-25)"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI;
                     Cloned unidirectionally. Primer: Oligo dT. Average insert
                     size 1.7 kb. Constructed by Life Technologies. Note: This
                     is a Xenopus Gene Collection (XGC) library."
BASE COUNT       20 a     20 c     24 g     30 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 94;
  Best Local Similarity   100.0%;  Pred. No. 3.5e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
            ||||||||
Db      72 AACGTTCG 79


RESULT  30
BE576515/c
LOCUS       BE576515       94 bp    mRNA            EST       15-AUG-2000
DEFINITION  dc40g03.y1 NICHD XGC Emb3 Xenopus laevis cDNA clone IMAGE:3399604
            5', mRNA sequence.
ACCESSION   BE576515
VERSION     BE576515.1  GI:9826314
KEYWORDS    EST.
SOURCE      African clawed frog.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 94)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Other_ESTs: dc40g03.x1
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Martha Rebbert, Steven L. Klein, Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: Xenopus clones from this library are available
            through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
            Seq primer: -40RP from Gibco
            High quality sequence stop: 83.
FEATURES             Location/Qualifiers
     source          1. .94
                     /organism="Xenopus laevis"
                     /db_xref="taxon:8355"
                     /clone="IMAGE:3399604"
                     /clone_lib="NICHD XGC Emb3"
                     /tissue_type="embryo (stages 24-25)"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI;
                     Cloned unidirectionally. Primer: Oligo dT. Average insert
                     size 1.7 kb. Constructed by Life Technologies. Note: This
                     is a Xenopus Gene Collection (XGC) library."
BASE COUNT       20 a     20 c     24 g     30 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 94;
  Best Local Similarity   100.0%;  Pred. No. 3.5e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
            ||||||||
Db      77 AACGTTCG 70


RESULT  31
AU013893/c
LOCUS       AU013893       98 bp    mRNA            EST       03-AUG-1998
DEFINITION  AU013893 Schizosaccharomyces pombe late log phase cDNA
            Schizosaccharomyces pombe cDNA clone spc08815, mRNA sequence.
ACCESSION   AU013893
VERSION     AU013893.1  GI:3368684
KEYWORDS    EST.
SOURCE      fission yeast.
  ORGANISM  Schizosaccharomyces pombe
            Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
            Schizosaccharomycetales; Schizosaccharomycetaceae;
            Schizosaccharomyces.
REFERENCE   1  (bases 1 to 98)
  AUTHORS   Morimyo,M. and Mita,K.
  TITLE     Identification of expressed sequence tags of Schizosaccharomyces
            pombe
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mitsuoki Morimyo
            Genome Research Group
            National Institute of Radiological Sciences
            9-1, Anagawa-4-chome, Inage-ku, Chiba, Chiba 263-8555, Japan
            Email: morimyo@nirs.go.jp.
FEATURES             Location/Qualifiers
     source          1. .98
                     /organism="Schizosaccharomyces pombe"
                     /strain="972"
                     /db_xref="taxon:4896"
                     /clone="spc08815"
                     /clone_lib="Schizosaccharomyces pombe late log phase cDNA"
                     /sex="h minus"
                     /note="Vector: M13mp19; The cDNA library of
                     Schizosaccharomyces pombe was prepared by cloning cDNA
                     into the SmaI site of M13mp19 DNA and the direction of DNA
                     sequences was not always from 5' to 3'. The cDNA data of
                     Schizosaccharomyces pombe are available for searching on
                     the World Wide Web.  (URL, http://www.nirs.go.jp)"
BASE COUNT       24 a     19 c     18 g     29 t      8 others
ORIGIN
```

```
  Query Match             100.0%;  Score 8;  DB 10;  Length 98;
  Best Local Similarity   100.0%;  Pred. No. 3.5e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      22 AACGTTCG 15


RESULT  32
AI105877
LOCUS       AI105877       99 bp    mRNA            EST       25-AUG-1998
DEFINITION  ab01f04_t3 ZF adult heart library Danio rerio cDNA 5 prime, mRNA
            sequence.
ACCESSION   AI105877
VERSION     AI105877.1  GI:3460980
KEYWORDS    EST.
SOURCE      zebrafish.
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Rasborinae; Danio.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Chen,J.N., DeSauvage,F., Hosobuchi,M., Jackson,D.G. and Fishman
            ,M.C.
  TITLE     Expressed Sequences from The Adult Zebrafish Heart
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark C. Fishman
            Cardiovascular Research Center
            Massachusetts General Hospital
            Mail code 1494100A, 149 13th Street, Charlestown, MA 02129, USA
            Fax: 6177265806
            Email: fishman@mgh.cvrc.harvard.edu
            http://zebrafish.mgh.harvard.edu
            The original clones used for sequencing are no longer available;
            the library is available from Mark C. Fishman.
            Insert Length: 99    Std Error: 0.00
            Seq primer: t3.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Danio rerio"
                     /strain="AB"
                     /db_xref="taxon:7955"
                     /clone_lib="ZF adult heart library"
                     /sex="mixed"
                     /tissue_type="myocardium, endocardium, vessel"
                     /dev_stage="adult"
                     /lab_host="E. coli XL1 Blue"
                     /note="Organ: heart; Vector: LambdaZAPII; Site_1: EcoR1;
                     Site_2: Xho1"
BASE COUNT       31 a     19 c     17 g     28 t      4 others
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 99;
  Best Local Similarity   100.0%;  Pred. No. 3.5e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db       9 AACGTTCG 16


RESULT  33
AW609278
LOCUS       AW609278       100 bp    mRNA            EST       23-MAR-2000
DEFINITION  MR3-ST0192-010200-206-g09 ST0192 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW609278
VERSION     AW609278.1  GI:7314019
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 100)
  AUTHORS   HCGP http://www.ludwig.org.br/ORESTES.
  TITLE     The FAPESP/LICR Human Cancer Genome Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?ti=MR3&t2=MR3-ST0192-
            010200-206-g09&t3=2000-02-01&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 97.
FEATURES             Location/Qualifiers
     source          1. .100
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="ST0192"
                     /dev_stage="Adult"
                     /note="Organ: stomach; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT       22 a     26 c     24 g     28 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 100;
  Best Local Similarity   100.0%;  Pred. No. 3.5e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTCG 8
           ||||||||
Db      16 AACGTTCG 23


RESULT  34
AA991491
LOCUS       AA991491       22 bp    mRNA            EST       03-JUN-1998
DEFINITION  os91h12.s1 NCI_CGAP_GC3 Homo sapiens cDNA clone IMAGE:1612775 3'
            similar to TR:O14597 O14597 NON-FUNCTIONAL FOLATE BINDING PROTEIN.
            ;, mRNA sequence.
ACCESSION   AA991491
VERSION     AA991491.1  GI:3177980
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
```

```
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1612775"
                     /clone_lib="NCI_CGAP_GC3"
                     /tissue_type="pooled germ cell tumors"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; 1st strand cDNA was prepared from 3 pooled
                     germ cell tumors, and was then primed with a Not I -
                     oligo(dT) primer. Double-stranded cDNA was ligated to Eco
                     RI adaptors (Pharmacia), digested with Not I and cloned
                     into the Not I and Eco RI sites of the modified pT7T3
                     vector.  Library is not normalized. Library was
                     constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT        4 a      4 c      9 g      5 t
ORIGIN


  Query Match             87.5%;  Score 7;  DB 10;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 ACGTTCG 8
           |||||||
Db       8 ACGTTCG 14


RESULT  35
AZ430042/c
LOCUS       AZ430042       22 bp    DNA             GSS       03-OCT-2000
DEFINITION  1M0214O13F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0214O13 F, DNA sequence.
ACCESSION   AZ430042
VERSION     AZ430042.1  GI:10554055
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0214  row: O  column: 13
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 22.
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0214O13"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT        4 a      6 c      3 g      9 t
ORIGIN


  Query Match             87.5%;  Score 7;  DB 13;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTC 7
           |||||||
Db      10 AACGTTC 4


RESULT  36
TA114E04P
LOCUS       TA114E04P      25 bp    DNA             GSS       13-DEC-2000
DEFINITION  T. brucei sheared genomic DNA clone 114e04, forward sequence,
            genomic survey sequence.
ACCESSION   AL462601
VERSION     AL462601.1  GI:11832406
KEYWORDS    GSS.
SOURCE      Trypanosoma brucei.
  ORGANISM  Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 25)
  AUTHORS   Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
            Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
            Melville,S.E., Rajandream,M.A. and Barrell,B.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
            project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
            Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
            nh1@sanger.ac.uk
COMMENT     Constructed at the Institute for Genomic Research (TIGR),
            Rockville, MD. Genomic DNA isolated from a cloned population of
            Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
            to give a tight size distribution (
            4 kb).  The v + i method used for the library construction is
            described in detail in Smith, H. and Venter, J.C. (Making small
            insert libraries for whole genome shotgun sequencing projects.  In
            Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
            Barrell, Oxford University Press, 1999).
            Email: nelsayed@tigr.org
            Details of T. brucei sequencing at the Sanger Centre are available
            at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES             Location/Qualifiers
     source          1. .25
                     /organism="Trypanosoma brucei"
                     /strain="TREU927"
                     /db_xref="taxon:5691"
```

```
                     /clone="114e04"
BASE COUNT         9 a      5 c      4 g      7 t
ORIGIN


  Query Match             87.5%;  Score 7;  DB 13;  Length 25;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 ACGTTCG 8
           |||||||
Db       4 ACGTTCG 10


RESULT  37
AZ823311
LOCUS       AZ823311       26 bp    DNA             GSS       20-FEB-2001
DEFINITION  2M0097H17F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0097H17 F, DNA sequence.
ACCESSION   AZ823311
VERSION     AZ823311.1  GI:12993219
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 26)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0097  row: H  column: 17
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 26.
FEATURES             Location/Qualifiers
     source          1. .26
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0097H17"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT         10 a      6 c      3 g      7 t
ORIGIN


  Query Match             87.5%;  Score 7;  DB 13;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTC 7
           |||||||
Db       4 AACGTTC 10


RESULT  38
TA69D08Q/c
LOCUS       TA69D08Q       26 bp    DNA             GSS       13-DEC-2000
DEFINITION  T. brucei sheared genomic DNA clone 69d08, reverse sequence,
            genomic survey sequence.
ACCESSION   AL458491
VERSION     AL458491.1  GI:11859115
KEYWORDS    GSS.
SOURCE      Trypanosoma brucei.
  ORGANISM  Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 26)
  AUTHORS   Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
            Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
            Melville,S.E., Rajandream,M.A. and Barrell,B.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
            project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
            Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
            nh1@sanger.ac.uk
COMMENT     Constructed at the Institute for Genomic Research (TIGR),
            Rockville, MD. Genomic DNA isolated from a cloned population of
            Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
            to give a tight size distribution (
            4 kb).  The v + i method used for the library construction is
            described in detail in Smith, H. and Venter, J.C. (Making small
            insert libraries for whole genome shotgun sequencing projects.  In
            Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
            Barrell, Oxford University Press, 1999).
            Email: nelsayed@tigr.org
            Details of T. brucei sequencing at the Sanger Centre are available
            at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES             Location/Qualifiers
     source          1. .26
                     /organism="Trypanosoma brucei"
                     /strain="TREU927"
                     /db_xref="taxon:5691"
                     /clone="69d08"
BASE COUNT         9 a      4 c     10 g      3 t
ORIGIN


  Query Match             87.5%;  Score 7;  DB 13;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 ACGTTCG 8
           |||||||
Db      22 ACGTTCG 16


RESULT  39
AZ974368
LOCUS       AZ974368       27 bp    DNA             GSS       27-APR-2001
DEFINITION  2M0248J21R Mouse 10kb plasmid UUGC2M library Mus musculus genomic
            clone UUGC2M0248J21 R, DNA sequence.
ACCESSION   AZ974368
```

```
VERSION     AZ974368.1  GI:13845595
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 27)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0248  row: J  column: 21
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 27.
FEATURES             Location/Qualifiers
     source          1. .27
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0248J21"
                     /clone_lib="Mouse 10kb plasmid UUGC2M library"
                     /sex="Female"
                     /lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (female) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT        6 a      12 c       4 g       5 t
ORIGIN


  Query Match             87.5%;  Score 7;  DB 13;  Length 27;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTC 7
           |||||||
Db       7 AACGTTC 13


RESULT  40
AA027602
LOCUS       AA027602        28 bp    mRNA            EST       11-SEP-1996
DEFINITION  mi12d08.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone
            IMAGE:463311 5' similar to SW:ARDH_HUMAN P41227 N-TERMINAL
            ACETYLTRANSFERASE COMPLEX ARD1 SUBUNIT HOMOLOG. ;, mRNA sequence.
ACCESSION   AA027602
VERSION     AA027602.1  GI:1493723
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 28)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:277127
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .28
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:463311"
                     /clone_lib="Soares mouse p3NMF19.5"
                     /dev_stage="19.5 dpc total fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Vector: pT7T3D (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by Bento
                     Soares and M.Fatima Bonaldo. RNA was kindly provided by
                     Dr. Minoru Ko (Wayne State University)."
BASE COUNT        8 a      8 c      6 g      6 t
ORIGIN


  Query Match             87.5%;  Score 7;  DB 10;  Length 28;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTC 7
           |||||||
Db      20 AACGTTC 26


RESULT  41
AI441029
LOCUS       AI441029        28 bp    mRNA            EST       01-DEC-1999
DEFINITION  sa58e02.y1 Gm-c1004 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
            Gm-c1004-3507 5' similar to TR:Q41454 Q41454 HMG-COA REDUCTASE ;,
            mRNA sequence.
ACCESSION   AI441029
VERSION     AI441029.1  GI:4286315
KEYWORDS    EST.
SOURCE      soybean.
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
```

Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE 1 (bases 1 to 28)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann ,R., Waterston,R. and Wilson,R.
TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomesystems.com web site: www.genomesystems.com
Possible reversed clone: similarity on wrong strand
Seq primer: -40RP from Gibco
High quality sequence stop: 1.

FEATURES Location/Qualifiers
source 1. .28
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1004-3507"
/clone_lib="Gm-c1004"
/tissue_type="root"
/lab_host="XL10-Gold"
/note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2: XhoI; Root cDNA. The mRNA was isolated from entire roots of 8 day old 'Williams' seedlings which were propagated on paper towels with distilled water. Stratagene's cDNA Synthesis Kit (catalog #200401) was used to synthesize the cDNA. First- strand synthesis was performed with 5-methyl dCTP, hence the ligated cDNA is hemimethylated. Stratagene's first-strand synthesis primer was used [GAGAGAGAGAGAGAGAGAACTAGTCTCGAG(T)-18]. After second-strand synthesis, the cDNA ends were 'polished' with clone Pfu DNA polymerase, ligated to EcoRI adapters, and phosphorylated. The XhoI site within the first-strand synthesis primer was restricted by digestion with XhoI; all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500bp cutoff, using GibcoBRL Life Technologies' cDNA Size Fractionation column. The column eluent was then ligated into Stratagene's pBluescript II XR Predigested vector (pBluescript II SK(+) that had been digested with EcoRI and XhoI, and phosphorylated). Both the white and blue colonies appear to contain recombinant plasmids with cDNA inserts. Blue colonies 9n=15) have been sequenced, and possess putative cDNA inserts. This library was constructed by Dr. Paul Keim & Virginia H. Coryell, Department of Biology, Box5640, Northern Arizona University, Flagstaff, AZ 86011, Phone: 520-523-1078 (Dr. Paul Keim), 520-523-1372 (Virginia H. Coryell), Fax: 520-523-7500, email: paul.keim@nau.edu, virginia.coryell@nau.edu"

BASE COUNT 4 a 12 c 7 g 5 t
ORIGIN

Query Match 87.5%; Score 7; DB 10; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.8e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ACGTTCG 8
|||||||
Db 3 ACGTTCG 9

RESULT 42
AA901420/c
LOCUS AA901420 31 bp mRNA EST 26-MAR-1999
DEFINITION SWOv3MCAM03G05 Onchocerca volvulus molting L3 larva cDNA (SL96MLW-OvmL3) Onchocerca volvulus cDNA clone onch77 5' similar to SW:CYTX_ONCVO P22085 ONCHOCYSTATIN PRECURSOR ;, mRNA sequence.
ACCESSION AA901420
VERSION AA901420.1 GI:3037174
KEYWORDS EST.
SOURCE Onchocerca volvulus.
ORGANISM Onchocerca volvulus
Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea; Onchocercidae; Onchocerca.
REFERENCE 1 (bases 1 to 31)
AUTHORS Williams,S.A., Lizotte-Waniewski,M., Laney,S., Lustigman,S., Hillier,L., Allen,M., Bowles,L., Geisel,S., Jost,S., Kucaba,T., Martin,J., Steptoe,M., Theising,B., White,Y., Wylie,T., Chappell,J. , Person,B., Gibbons,M., Harvey,N., Pape,D., Chamberlain,A., Morales,R., Schurk,R., Ritter,E., Kohn,S., Underwood,K. and Marra ,M.
TITLE Molecular Parasitology OvmL3
JOURNAL Unpublished (1998)
COMMENT Contact: Steven A. Williams
Molecular Parasitology
Smith College Department of Biological Sciences
Department of Biological Sciences, Clark Science Center, Smith College, Northampton, MA, 01063, USA
Tel: 4135853826
Fax: 4135853786
Email: genome@smith.edu
The library was constructed by Sara Lustigman and Michelle Lizotte-Waniewski in the Laboratory of Dr. S.A.Williams. The library is available from Dr. Sara Lustigman email slustigm@nybc.org When requesting this clone from Dr. Lustigman, please reference the Williams lab clone id - SWOv3MCAM03G05
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 1.

FEATURES Location/Qualifiers
source 1. .31
/organism="Onchocerca volvulus"
/strain="Kumba, Cameroons"
/db_xref="taxon:6282"
/clone="onch77"
/clone_lib="Onchocerca volvulus molting L3 larva cDNA (SL96MLW-OvmL3)"
/dev_stage="molting L3"
/lab_host="XL1-Blue MRF'"
/note="Vector: Lambda Uni-ZAP XR; Site_1: Eco RI; Site_2: Xho I; Filarial nematode parasite of humans. Third-stage larvae, L3, were isolated from infected black flies in Cameroon (forest strain). The L3 were cultured in 20% FCS in IMDM+ NCTC 135 and collected after day 1, 2, or 3 in culture. L3 of O. volvulus molt to fourth-stage larvae by day 5 in culture. mRNA was isolated from approximately 6000 molting larvae (mL3),2000 larvae from day 1, 2 or 3 in culture, and converted to double-stranded cDNA using reverse transcriptase and oligo(dT) followed by RNase H and DNA pol I. The library was constructed in the lambda Uni-Zap XR vector and has 1 x 10E6 independent recombinants and the average insert size is ~1200 bp. The library was constructed by Sara Lustigman and Michelle Lizotte-Waniewski in the Laboratory of Dr. S. A. Williams. The library is available from Dr. Sara Lustigman (email: slustigm@nybc.org)."

BASE COUNT 10 a 3 c 10 g 8 t
ORIGIN

Query Match 87.5%; Score 7; DB 10; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;

```
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTC 7
           |||||||
Db      24 AACGTTC 18


RESULT  43
AA910190
LOCUS       AA910190       31 bp    mRNA            EST       09-JUN-1998
DEFINITION  oj29a03.s1 NCI_CGAP_Kid3 Homo sapiens cDNA clone IMAGE:1493548 3'
            similar to TR:Q14291 Q14291 FOCAL ADHESION KINASE. ;, mRNA
            sequence.
ACCESSION   AA910190
VERSION     AA910190.1  GI:3049480
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 31)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Insert Length: 800   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .31
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1493548"
                     /clone_lib="NCI_CGAP_Kid3"
                     /lab_host="DH10B"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer,
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not
                     I and Eco RI sites  of the modified pT7T3 vector.  mRNA
                     source: 2 pooled kidneys. Library went through one round
                     of normalization. Library constructed by Bento Soares and
                     M. Fatima Bonaldo.  "
BASE COUNT        7 a     13 c      6 g      5 t
ORIGIN


  Query Match             87.5%;  Score 7;  DB 10;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AACGTTC 7
           |||||||
Db      18 AACGTTC 24


RESULT  44
AA910190/c
LOCUS       AA910190       31 bp    mRNA            EST       09-JUN-1998
DEFINITION  oj29a03.s1 NCI_CGAP_Kid3 Homo sapiens cDNA clone IMAGE:1493548 3'
            similar to TR:Q14291 Q14291 FOCAL ADHESION KINASE. ;, mRNA
            sequence.
ACCESSION   AA910190
VERSION     AA910190.1  GI:3049480
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 31)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Insert Length: 800   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .31
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1493548"
                     /clone_lib="NCI_CGAP_Kid3"
                     /lab_host="DH10B"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer,
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not
                     I and Eco RI sites  of the modified pT7T3 vector.  mRNA
                     source: 2 pooled kidneys. Library went through one round
                     of normalization. Library constructed by Bento Soares and
                     M. Fatima Bonaldo.  "
BASE COUNT        7 a     13 c      6 g      5 t
ORIGIN


  Query Match             87.5%;  Score 7;  DB 10;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       1 AACGTTC 7
           |||||||
Db      23 AACGTTC 17


RESULT  45
AI002314
LOCUS       AI002314        34 bp    mRNA            EST       09-JUN-1998
DEFINITION  or73c09.s1 NCI_CGAP_Lu5 Homo sapiens cDNA clone IMAGE:1601488 3'
            similar to TR:Q15733 Q15733 PHOSPHATIDYLINOSITOL ;, mRNA sequence.
ACCESSION   AI002314
VERSION     AI002314.1  GI:3202648
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 34)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .34
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1601488"
                     /clone_lib="NCI_CGAP_Lu5"
                     /tissue_type="carcinoid"
                     /lab_host="DH10B"
                     /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; 1st strand cDNA was prepared from
                     neuroendocrine lung carcinoid, and was then primed with a
                     Not I - oligo(dT) primer. Double-stranded cDNA was ligated
                     to Eco RI adaptors (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of the modified
                     pT7T3 vector.  Library is normalized. Library was
                     constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT        7 a      12 c       7 g       8 t
ORIGIN


  Query Match             87.5%;  Score 7;  DB 10;  Length 34;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 ACGTTCG 8
           |||||||
Db      23 ACGTTCG 29


Search completed: November 29, 2001, 14:23:46
Job time: 8079 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: November 29, 2001, 14:47:06 ; Search time 1391.6 Seconds
(without alignments)
94.839 Million cell updates/sec

Title: FRAG2
Perfect score: 8
Sequence: 1 GACGTTCG 8

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 661134

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vi:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rod:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rod:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	Query Match	Length	DB	ID	Description
c	1	8	100.0	15	6	A89447	A89447 Sequence 15
c	2	8	100.0	16	6	A89446	A89446 Sequence 15
c	3	8	100.0	17	6	A60700	A60700 Sequence 8
c	4	8	100.0	17	6	AR125077	AR125077 Sequence
	5	8	100.0	17	6	AX139244	AX139244 Sequence
	6	8	100.0	18	6	AX028711	AX028711 Sequence
	7	8	100.0	19	6	AX132650	AX132650 Sequence
	8	8	100.0	19	6	AX132651	AX132651 Sequence
	9	8	100.0	20	6	A89786	A89786 Sequence 8
	10	8	100.0	20	6	A90873	A90873 Sequence 8
c	11	8	100.0	20	6	AR096949	AR096949 Sequence
	12	8	100.0	20	6	AR100579	AR100579 Sequence
	13	8	100.0	20	6	AR100585	AR100585 Sequence
c	14	8	100.0	20	6	AR107432	AR107432 Sequence
c	15	8	100.0	20	6	AR156714	AR156714 Sequence
	16	8	100.0	20	6	AX104329	AX104329 Sequence
	17	8	100.0	20	6	AX104332	AX104332 Sequence
	18	8	100.0	20	6	AX104334	AX104334 Sequence
	19	8	100.0	20	6	AX104664	AX104664 Sequence
	20	8	100.0	20	6	AX104705	AX104705 Sequence
c	21	8	100.0	20	6	I43022	I43022 Sequence 4
c	22	8	100.0	20	6	I49077	I49077 Sequence 10
	23	8	100.0	21	6	AR100580	AR100580 Sequence
	24	8	100.0	21	6	AR100586	AR100586 Sequence
c	25	8	100.0	21	6	AX004662	AX004662 Sequence
	26	8	100.0	22	6	AX104693	AX104693 Sequence
	27	8	100.0	22	6	AX104789	AX104789 Sequence
	28	8	100.0	22	6	AX104846	AX104846 Sequence
	29	8	100.0	22	6	AX105122	AX105122 Sequence
	30	8	100.0	22	6	AX105255	AX105255 Sequence
	31	8	100.0	23	6	AR096950	AR096950 Sequence
c	32	8	100.0	23	6	AR137715	AR137715 Sequence
c	33	8	100.0	23	6	AR137722	AR137722 Sequence
	34	8	100.0	23	6	I43023	I43023 Sequence 5
	35	8	100.0	24	6	I38780	I38780 Sequence 18
	36	8	100.0	26	6	AR099214	AR099214 Sequence
	37	8	100.0	26	6	AR154308	AR154308 Sequence
c	38	8	100.0	26	6	I49791	I49791 Sequence 14
c	39	8	100.0	27	6	AR121792	AR121792 Sequence
	40	8	100.0	31	6	AR001148	AR001148 Sequence
	41	8	100.0	31	6	AR003026	AR003026 Sequence
	42	8	100.0	31	6	AR033000	AR033000 Sequence
	43	8	100.0	31	6	AR126490	AR126490 Sequence
	44	8	100.0	31	6	AR126494	AR126494 Sequence
	45	8	100.0	31	6	I76870	I76870 Sequence 12

ALIGNMENTS

RESULT 1
A89447/c
LOCUS A89447 15 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 1595 from Patent WO9833904.
ACCESSION A89447
VERSION A89447.1 GI:6738017
KEYWORDS .
SOURCE unidentified.
ORGANISM unidentified
unclassified.
REFERENCE 1 (bases 1 to 15)
AUTHORS Brysch,W. and Schlingensiepen,K.
TITLE AN ANTISENSE OLIGONUCLEOTIDE PREPARATION METHOD
JOURNAL Patent: WO 9833904-A 1595 06-AUG-1998;
BIOGNOSTIK GES (DE); BRYSCH WOLFGANG (DE)
FEATURES Location/Qualifiers
source 1. .15
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 2 a 4 c 5 g 4 t

```
ORIGIN


  Query Match             100.0%;  Score 8;  DB 6;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db     10 GACGTTCG 3


RESULT   2
A89446/c
LOCUS       A89446          16 bp    DNA              PAT       22-JAN-2000
DEFINITION  Sequence 1594 from Patent WO9833904.
ACCESSION   A89446
VERSION     A89446.1  GI:6738016
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 16)
  AUTHORS   Brysch,W. and Schlingensiepen,K.
  TITLE     AN ANTISENSE OLIGONUCLEOTIDE PREPARATION METHOD
  JOURNAL   Patent: WO 9833904-A 1594 06-AUG-1998;
            BIOGNOSTIK GES (DE); BRYSCH WOLFGANG (DE)
FEATURES             Location/Qualifiers
     source          1. .16
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT         2 a      4 c      5 g      5 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 6;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db      8 GACGTTCG 1


RESULT   3
A60700/c
LOCUS       A60700          17 bp    DNA              PAT       06-MAR-1998
DEFINITION  Sequence 8 from Patent WO9708320.
ACCESSION   A60700
VERSION     A60700.1  GI:3715348
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 17)
  AUTHORS   Knappik,A., Pack,P., Ilag,V., Ge,L., Moroney,S. and Plueckthun,A.
  TITLE     PROTEIN/(POLY)PEPTIDE LIBRARIES
  JOURNAL   Patent: WO 9708320-A 8 06-MAR-1997;
            MORPHOSYS PROTEINOPTIMIERUNG (DE)
FEATURES             Location/Qualifiers
     source          1. .17
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT         5 a      5 c      6 g      1 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 6;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db     16 GACGTTCG 9


RESULT   4
AR125077/c
LOCUS       AR125077        17 bp    DNA              PAT       16-MAY-2001
DEFINITION  Sequence 18 from patent US 6177075.
ACCESSION   AR125077
VERSION     AR125077.1  GI:14111139
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 17)
  AUTHORS   Christian,P.Daniel, Gordon,K.HienrichJulius and Hanzlik,T.Nelson.
  TITLE     Insect viruses and their uses in protecting plants
  JOURNAL   Patent: US 6177075-A 18 23-JAN-2001;
FEATURES             Location/Qualifiers
     source          1. .17
                     /organism="unknown"
BASE COUNT         6 a      3 c      6 g      2 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 6;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db     12 GACGTTCG 5


RESULT   5
AX139244
LOCUS       AX139244        17 bp    DNA              PAT       30-MAY-2001
DEFINITION  Sequence 92 from Patent EP1076099.
ACCESSION   AX139244
VERSION     AX139244.1  GI:14274917
KEYWORDS    .
SOURCE      Mycobacterium tuberculosis.
  ORGANISM  Mycobacterium tuberculosis
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Mycobacteriaceae;
            Mycobacterium; Mycobacterium tuberculosis complex.
REFERENCE   1  (bases 1 to 17)
  AUTHORS   Suzuki,Y., Nishida,M. and Takenishi,S.
  TITLE     Kit for diagnosis of tubercle bacilli
  JOURNAL   Patent: EP 1076099-A 92 14-FEB-2001;
            NISSHINBO INDUSTRIES, INC. (JP) ; System Research Incorporation
            (JP)
FEATURES             Location/Qualifiers
     source          1. .17
                     /organism="Mycobacterium tuberculosis"
                     /db_xref="taxon:1773"
                     /note="capture"
BASE COUNT         3 a      4 c      5 g      5 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 6;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db      7 GACGTTCG 14


RESULT   6
AX028711
LOCUS       AX028711        18 bp    DNA              PAT       16-SEP-2000
DEFINITION  Sequence 15 from Patent EP1018550.
```

```
ACCESSION   AX028711
VERSION     AX028711.1  GI:10189824
KEYWORDS    .
SOURCE      European house dust mite.
  ORGANISM  Dermatophagoides pteronyssinus
            Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
            Acariformes; Sarcoptiformes; Astigmata; Analgoidea; Pyroglyphidae;
            Dermatophagoides.
REFERENCE   1  (bases 1 to 18)
  AUTHORS   Thomas,W.R. and Chua,K.Y.
  TITLE     Allergenic protein and peptides from house dust mite and uses
            thereof
  JOURNAL   Patent: EP 1018550-A 15 12-JUL-2000;
            INST CHILD HEALTH RESEARCH (AU)
FEATURES             Location/Qualifiers
     source          1. .18
                     /organism="Dermatophagoides pteronyssinus"
                     /db_xref="taxon:6956"
BASE COUNT        5 a      3 c      3 g      7 t
ORIGIN


  Query Match           100.0%;  Score 8;  DB 6;  Length 18;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db      3 GACGTTCG 10


RESULT   7
AX132650
LOCUS       AX132650       19 bp    DNA             PAT       15-MAY-2001
DEFINITION  Sequence 3868 from Patent WO0130362.
ACCESSION   AX132650
VERSION     AX132650.1  GI:14138955
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 19)
  AUTHORS   Robbins,J.M. and Tritz,R.
  TITLE     Ribozyme therapy for the treatment of proliferative skin and eye
            diseases
  JOURNAL   Patent: WO 0130362-A 3868 03-MAY-2001;
            IMMUSOL, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .19
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="PCNA HH ribozyme binding site"
BASE COUNT        1 a      9 c      5 g      4 t
ORIGIN


  Query Match           100.0%;  Score 8;  DB 6;  Length 19;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db      5 GACGTTCG 12


RESULT   8
AX132651
LOCUS       AX132651       19 bp    DNA             PAT       15-MAY-2001
DEFINITION  Sequence 3869 from Patent WO0130362.
ACCESSION   AX132651
VERSION     AX132651.1  GI:14138956
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 19)
  AUTHORS   Robbins,J.M. and Tritz,R.
  TITLE     Ribozyme therapy for the treatment of proliferative skin and eye
            diseases
  JOURNAL   Patent: WO 0130362-A 3869 03-MAY-2001;
            IMMUSOL, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .19
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="PCNA HH ribozyme binding site"
BASE COUNT        1 a      9 c      5 g      4 t
ORIGIN


  Query Match           100.0%;  Score 8;  DB 6;  Length 19;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db      4 GACGTTCG 11


RESULT   9
A89786
LOCUS       A89786         20 bp    DNA             PAT       22-JAN-2000
DEFINITION  Sequence 8 from Patent WO9832462.
ACCESSION   A89786
VERSION     A89786.1  GI:6738300
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Lipford,G.B. and Heeg,K.
  TITLE     PHARMACEUTICAL COMPOSITIONS COMPRISING A POLYNUCLEOTIDE AND
            OPTIONALLY AN ANTIGEN ESPECIALLY FOR VACCINATION
  JOURNAL   Patent: WO 9832462-A 8 30-JUL-1998;
            LIPFORD GRAYSON B (DE); HEEG KLAUS (DE)
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT        4 a      5 c      6 g      5 t
ORIGIN


  Query Match           100.0%;  Score 8;  DB 6;  Length 20;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db      8 GACGTTCG 15


RESULT  10
A90873
LOCUS       A90873         20 bp    DNA             PAT       22-JAN-2000
DEFINITION  Sequence 8 from Patent EP0855184.
ACCESSION   A90873
VERSION     A90873.1  GI:6739267
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Heeg,K.P. and Lipford,G.B.
```

```
  TITLE     Pharmaceutical composition comprising a polynucleotide and an
            antigen especially for vaccination
  JOURNAL   Patent: EP 0855184-A 8 29-JUL-1998;
            HEEG KLAUS PROF DR (DE); LIPFORD GRAYSON B DR (DE)
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT        4 a      5 c      6 g      5 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 6;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db       8 GACGTTCG 15


RESULT  11
AR096949/c
LOCUS       AR096949       20 bp    DNA             PAT       14-FEB-2001
DEFINITION  Sequence 4 from patent US 6071480.
ACCESSION   AR096949
VERSION     AR096949.1  GI:12805679
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Halaka,F.G.
  TITLE     Method for generating a standing sonic wave, methods of sonication
            with a standing sonic wave, and a standing sonic wave sonicator
  JOURNAL   Patent: US 6071480-A 4 06-JUN-2000;
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="unknown"
BASE COUNT        6 a      5 c      6 g      3 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 6;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      18 GACGTTCG 11


RESULT  12
AR100579
LOCUS       AR100579       20 bp    DNA             PAT       14-FEB-2001
DEFINITION  Sequence 95 from patent US 6080588.
ACCESSION   AR100579
VERSION     AR100579.1  GI:12811027
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Glick,G.D.
  TITLE     Therapeutic methods for benzodiazepine derivatives
  JOURNAL   Patent: US 6080588-A 95 27-JUN-2000;
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="unknown"
BASE COUNT        2 a      4 c      8 g      6 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 6;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      11 GACGTTCG 18


RESULT  13
AR100585
LOCUS       AR100585       20 bp    DNA             PAT       14-FEB-2001
DEFINITION  Sequence 103 from patent US 6080588.
ACCESSION   AR100585
VERSION     AR100585.1  GI:12811033
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Glick,G.D.
  TITLE     Therapeutic methods for benzodiazepine derivatives
  JOURNAL   Patent: US 6080588-A 103 27-JUN-2000;
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="unknown"
BASE COUNT        2 a      4 c      8 g      6 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 6;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      11 GACGTTCG 18


RESULT  14
AR107432/c
LOCUS       AR107432       20 bp    DNA             PAT       14-FEB-2001
DEFINITION  Sequence 15 from patent US 6110464.
ACCESSION   AR107432
VERSION     AR107432.1  GI:12822919
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Malvar,T. and Gilmer,A.Jelen.
  TITLE     Broad-spectrum .delta.-endotoxins
  JOURNAL   Patent: US 6110464-A 15 29-AUG-2000;
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="unknown"
BASE COUNT        6 a      6 c      2 g      6 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 6;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      17 GACGTTCG 10


RESULT  15
AR156714/c
LOCUS       AR156714       20 bp    DNA             PAT       08-AUG-2001
DEFINITION  Sequence 15 from patent US 6242241.
```

```
ACCESSION   AR156714
VERSION     AR156714.1  GI:15125418
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Malvar,T. and Gilmer,A.Jelen.
  TITLE     Polynucleotide compositions encoding broad-spectrum
            .delta.-endotoxins
  JOURNAL   Patent: US 6242241-A 15 05-JUN-2001;
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="unknown"
BASE COUNT        6 a      6 c      2 g      6 t
ORIGIN

  Query Match             100.0%;  Score 8;  DB 6;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      17 GACGTTCG 10


RESULT  16
AX104329
LOCUS       AX104329       20 bp    DNA             PAT       30-APR-2001
DEFINITION  Sequence 521 from Patent WO0122972.
ACCESSION   AX104329
VERSION     AX104329.1  GI:13920526
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Krieg,A.M., Schetter,C. and Vollmer,J.C.
  TITLE     Immunostimulatory nucleic acids
  JOURNAL   Patent: WO 0122972-A 521 05-APR-2001;
            UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical
            GmbH (DE)
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT        2 a      3 c     12 g      3 t
ORIGIN

  Query Match             100.0%;  Score 8;  DB 6;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db       8 GACGTTCG 15


RESULT  17
AX104332
LOCUS       AX104332       20 bp    DNA             PAT       30-APR-2001
DEFINITION  Sequence 524 from Patent WO0122972.
ACCESSION   AX104332
VERSION     AX104332.1  GI:13920529
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Krieg,A.M., Schetter,C. and Vollmer,J.C.
  TITLE     Immunostimulatory nucleic acids
  JOURNAL   Patent: WO 0122972-A 524 05-APR-2001;
            UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical
            GmbH (DE)
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT        2 a      3 c     12 g      3 t
ORIGIN

  Query Match             100.0%;  Score 8;  DB 6;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db       8 GACGTTCG 15


RESULT  18
AX104334
LOCUS       AX104334       20 bp    DNA             PAT       30-APR-2001
DEFINITION  Sequence 526 from Patent WO0122972.
ACCESSION   AX104334
VERSION     AX104334.1  GI:13920531
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Krieg,A.M., Schetter,C. and Vollmer,J.C.
  TITLE     Immunostimulatory nucleic acids
  JOURNAL   Patent: WO 0122972-A 526 05-APR-2001;
            UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical
            GmbH (DE)
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT        6 a      3 c      8 g      3 t
ORIGIN

  Query Match             100.0%;  Score 8;  DB 6;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db       8 GACGTTCG 15


RESULT  19
AX104664
LOCUS       AX104664       20 bp    DNA             PAT       30-APR-2001
DEFINITION  Sequence 856 from Patent WO0122972.
ACCESSION   AX104664
VERSION     AX104664.1  GI:13920861
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Krieg,A.M., Schetter,C. and Vollmer,J.C.
  TITLE     Immunostimulatory nucleic acids
  JOURNAL   Patent: WO 0122972-A 856 05-APR-2001;
            UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical
            GmbH (DE)
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
```

```
BASE COUNT        6 a      3 c      8 g      3 t
ORIGIN

  Query Match             100.0%;  Score 8;  DB 6;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db      8 GACGTTCG 15


RESULT  20
AX104705
LOCUS       AX104705       20 bp    DNA             PAT       30-APR-2001
DEFINITION  Sequence 897 from Patent WO0122972.
ACCESSION   AX104705
VERSION     AX104705.1  GI:13920902
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Krieg,A.M., Schetter,C. and Vollmer,J.C.
  TITLE     Immunostimulatory nucleic acids
  JOURNAL   Patent: WO 0122972-A 897 05-APR-2001;
            UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical
            GmbH (DE)
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT        1 a      9 c      8 g      2 t
ORIGIN

  Query Match             100.0%;  Score 8;  DB 6;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db      3 GACGTTCG 10


RESULT  21
I43022/c
LOCUS       I43022         20 bp    DNA             PAT       07-OCT-1997
DEFINITION  Sequence 4 from patent US 5631130.
ACCESSION   I43022
VERSION     I43022.1  GI:2468266
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Leckie,G.W., Davis,A.H., Semple-Facey,I.E., Manlove,M.T. and
            Solomon,N.A.
  TITLE     Materials and methods for the detection of Mycobacterium
            tuberculosis
  JOURNAL   Patent: US 5631130-A 4 20-MAY-1997;
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="unknown"
BASE COUNT        6 a      5 c      6 g      3 t
ORIGIN

  Query Match             100.0%;  Score 8;  DB 6;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db     18 GACGTTCG 11


RESULT  22
I49077/c
LOCUS       I49077         20 bp    DNA             PAT       07-OCT-1997
DEFINITION  Sequence 10 from patent US 5627195.
ACCESSION   I49077
VERSION     I49077.1  GI:2467540
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Hu,S.
  TITLE     Treatment for ocular inflammation
  JOURNAL   Patent: US 5627195-A 10 06-MAY-1997;
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="unknown"
BASE COUNT        4 a      7 c      4 g      5 t
ORIGIN

  Query Match             100.0%;  Score 8;  DB 6;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db     16 GACGTTCG 9


RESULT  23
AR100580
LOCUS       AR100580       21 bp    DNA             PAT       14-FEB-2001
DEFINITION  Sequence 96 from patent US 6080588.
ACCESSION   AR100580
VERSION     AR100580.1  GI:12811028
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Glick,G.D.
  TITLE     Therapeutic methods for benzodiazepine derivatives
  JOURNAL   Patent: US 6080588-A 96 27-JUN-2000;
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="unknown"
BASE COUNT        3 a      6 c      7 g      5 t
ORIGIN

  Query Match             100.0%;  Score 8;  DB 6;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db     12 GACGTTCG 19


RESULT  24
AR100586
LOCUS       AR100586       21 bp    DNA             PAT       14-FEB-2001
DEFINITION  Sequence 104 from patent US 6080588.
ACCESSION   AR100586
VERSION     AR100586.1  GI:12811034
KEYWORDS    .
SOURCE      Unknown.
```

```
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Glick,G.D.
  TITLE     Therapeutic methods for benzodiazepine derivatives
  JOURNAL   Patent: US 6080588-A 104 27-JUN-2000;
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="unknown"
BASE COUNT        3 a      5 c      6 g      7 t
ORIGIN


  Query Match           100.0%;  Score 8;  DB 6;  Length 21;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      12 GACGTTCG 19


RESULT  25
AX004662/c
LOCUS       AX004662       21 bp    DNA             PAT       24-AUG-2000
DEFINITION  Sequence 11 from Patent WO9915644.
ACCESSION   AX004662
VERSION     AX004662.1  GI:9928098
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Cardinal,G. and Levesque,R.C.
  TITLE     Method for the identification of essential genes and therapeutic
            targets
  JOURNAL   Patent: WO 9915644-A 11 01-APR-1999;
            CARDINAL GUY (CA); UNIV LAVAL (CA)
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="OLIGONUCLEOTIDE"
BASE COUNT        4 a      7 c      6 g      4 t
ORIGIN


  Query Match           100.0%;  Score 8;  DB 6;  Length 21;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      17 GACGTTCG 10


RESULT  26
AX104693
LOCUS       AX104693       22 bp    DNA             PAT       30-APR-2001
DEFINITION  Sequence 885 from Patent WO0122972.
ACCESSION   AX104693
VERSION     AX104693.1  GI:13920890
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Krieg,A.M., Schetter,C. and Vollmer,J.C.
  TITLE     Immunostimulatory nucleic acids
  JOURNAL   Patent: WO 0122972-A 885 05-APR-2001;
            UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical
            GmbH (DE)
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT        1 a      10 c      7 g      4 t
ORIGIN


  Query Match           100.0%;  Score 8;  DB 6;  Length 22;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db       5 GACGTTCG 12


RESULT  27
AX104789
LOCUS       AX104789       22 bp    DNA             PAT       30-APR-2001
DEFINITION  Sequence 981 from Patent WO0122972.
ACCESSION   AX104789
VERSION     AX104789.1  GI:13920986
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Krieg,A.M., Schetter,C. and Vollmer,J.C.
  TITLE     Immunostimulatory nucleic acids
  JOURNAL   Patent: WO 0122972-A 981 05-APR-2001;
            UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical
            GmbH (DE)
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT        3 a      3 c      13 g      3 t
ORIGIN


  Query Match           100.0%;  Score 8;  DB 6;  Length 22;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db       4 GACGTTCG 11


RESULT  28
AX104846
LOCUS       AX104846       22 bp    DNA             PAT       30-APR-2001
DEFINITION  Sequence 1038 from Patent WO0122972.
ACCESSION   AX104846
VERSION     AX104846.1  GI:13921043
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Krieg,A.M., Schetter,C. and Vollmer,J.C.
  TITLE     Immunostimulatory nucleic acids
  JOURNAL   Patent: WO 0122972-A 1038 05-APR-2001;
            UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical
            GmbH (DE)
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT        3 a      3 c      13 g      3 t
ORIGIN
```

```
  Query Match           100.0%;  Score 8;  DB 6;  Length 22;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db      4 GACGTTCG 11


RESULT 29
AX105122
LOCUS       AX105122        22 bp    DNA             PAT       30-APR-2001
DEFINITION  Sequence 20 from Patent WO0122990.
ACCESSION   AX105122
VERSION     AX105122.1  GI:13921272
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Hartmann,G.D., Bratzler,R.L. and Krieg,A.U.
  TITLE     Methods related to immunostimulatory nucleic acid-induced
            interferon
  JOURNAL   Patent: WO 0122990-A 20 05-APR-2001;
            Coley Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH
            FOUNDATION (US)
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Synthetic Oligonucleotide"
     misc_feature    1. .2
                     /note="Backbone has phosphorothioate linkages."
     misc_feature    3. .16
                     /note="Backbone has phosphodiester linkages."
     misc_feature    17. .21
                     /note="Backbone has phosphorothioate linkages."
     misc_feature    22
                     /note="Backbone has phosphodiester linkages."
BASE COUNT        3 a      3 c      13 g      3 t
ORIGIN


  Query Match           100.0%;  Score 8;  DB 6;  Length 22;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db      4 GACGTTCG 11


RESULT 30
AX105255
LOCUS       AX105255        22 bp    DNA             PAT       30-APR-2001
DEFINITION  Sequence 154 from Patent WO0122990.
ACCESSION   AX105255
VERSION     AX105255.1  GI:13921405
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Hartmann,G.D., Bratzler,R.L. and Krieg,A.U.
  TITLE     Methods related to immunostimulatory nucleic acid-induced
            interferon
  JOURNAL   Patent: WO 0122990-A 154 05-APR-2001;
            Coley Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH
            FOUNDATION (US)
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Synthetic Oligonucleotide"
     misc_feature    1. .22
                     /note="Backbone has phosphorothioate linkages."
BASE COUNT        3 a      3 c      13 g      3 t
ORIGIN


  Query Match           100.0%;  Score 8;  DB 6;  Length 22;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db      4 GACGTTCG 11


RESULT 31
AR096950
LOCUS       AR096950        23 bp    DNA             PAT       14-FEB-2001
DEFINITION  Sequence 5 from patent US 6071480.
ACCESSION   AR096950
VERSION     AR096950.1  GI:12805680
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 23)
  AUTHORS   Halaka,F.G.
  TITLE     Method for generating a standing sonic wave, methods of sonication
            with a standing sonic wave, and a standing sonic wave sonicator
  JOURNAL   Patent: US 6071480-A 5 06-JUN-2000;
FEATURES             Location/Qualifiers
     source          1. .23
                     /organism="unknown"
BASE COUNT        3 a      6 c      7 g      7 t
ORIGIN


  Query Match           100.0%;  Score 8;  DB 6;  Length 23;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db      3 GACGTTCG 10


RESULT 32
AR137715/c
LOCUS       AR137715        23 bp    DNA             PAT       16-JUN-2001
DEFINITION  Sequence 3 from patent US 6197556.
ACCESSION   AR137715
VERSION     AR137715.1  GI:14479224
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 23)
  AUTHORS   Ulanovsky,L. and Raja,M.C.
  TITLE     Nucleic acid amplification using modular branched primers
  JOURNAL   Patent: US 6197556-A 3 06-MAR-2001;
FEATURES             Location/Qualifiers
     source          1. .23
                     /organism="unknown"
BASE COUNT        5 a      8 c      5 g      3 t      2 others
ORIGIN


  Query Match           100.0%;  Score 8;  DB 6;  Length 23;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
```

```
          ||||||||
Db      16 GACGTTCG 9


RESULT  33
AR137722/c
LOCUS       AR137722        23 bp    DNA             PAT       16-JUN-2001
DEFINITION  Sequence 10 from patent US 6197556.
ACCESSION   AR137722
VERSION     AR137722.1  GI:14479231
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 23)
  AUTHORS   Ulanovsky,L. and Raja,M.C.
  TITLE     Nucleic acid amplification using modular branched primers
  JOURNAL   Patent: US 6197556-A 10 06-MAR-2001;
FEATURES             Location/Qualifiers
     source          1. .23
                     /organism="unknown"
BASE COUNT         3 a      6 c      8 g      4 t      2 others
ORIGIN


  Query Match             100.0%;  Score 8;  DB 6;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db      16 GACGTTCG 9


RESULT  34
I43023
LOCUS       I43023          23 bp    DNA             PAT       07-OCT-1997
DEFINITION  Sequence 5 from patent US 5631130.
ACCESSION   I43023
VERSION     I43023.1  GI:2468267
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 23)
  AUTHORS   Leckie,G.W., Davis,A.H., Semple-Facey,I.E., Manlove,M.T. and
            Solomon,N.A.
  TITLE     Materials and methods for the detection of Mycobacterium
            tuberculosis
  JOURNAL   Patent: US 5631130-A 5 20-MAY-1997;
FEATURES             Location/Qualifiers
     source          1. .23
                     /organism="unknown"
BASE COUNT         3 a      6 c      7 g      7 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 6;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db      3 GACGTTCG 10


RESULT  35
I38780
LOCUS       I38780          24 bp    DNA             PAT       13-MAY-1997
DEFINITION  Sequence 18 from patent US 5616461.
ACCESSION   I38780
VERSION     I38780.1  GI:2083258
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   Schaffer,P.A. and Dabrowski Amaral,C.E.
  TITLE     Assay for antiviral activity using complex of herpesvirus origin of
            replication and cellular protein
  JOURNAL   Patent: US 5616461-A 18 01-APR-1997;
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="unknown"
BASE COUNT         4 a      9 c      5 g      6 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 6;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db      3 GACGTTCG 10


RESULT  36
AR099214
LOCUS       AR099214        26 bp    DNA             PAT       14-FEB-2001
DEFINITION  Sequence 108 from patent US 6077692.
ACCESSION   AR099214
VERSION     AR099214.1  GI:12808980
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 26)
  AUTHORS   Ruben,S.M., Jimenez,P., Duan,D.Roxanne, Rampy,M.A., Mendrick,D.,
            Zhang,J., Ni,J., Moore,P.A., Coleman,T.A., Gruber,J.R., Dillon,P.J.
            and Gentz,R.L.
  TITLE     Keratinocyte growth factor-2
  JOURNAL   Patent: US 6077692-A 108 20-JUN-2000;
FEATURES             Location/Qualifiers
     source          1. .26
                     /organism="unknown"
BASE COUNT         2 a      6 c      12 g      6 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 6;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db      15 GACGTTCG 22


RESULT  37
AR154308
LOCUS       AR154308        26 bp    DNA             PAT       08-AUG-2001
DEFINITION  Sequence 29 from patent US 6238888.
ACCESSION   AR154308
VERSION     AR154308.1  GI:15122361
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 26)
  AUTHORS   Gentz,R.L., Chopra,A., Kaushal,P., Spitznagel,T., Unsworth,E. and
            Khan,F.
  TITLE     Keratinocyte growth factor-2 formulations
  JOURNAL   Patent: US 6238888-A 29 29-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .26
```

```
                     /organism="unknown"
BASE COUNT         2 a      6 c     12 g      6 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 6;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      15 GACGTTCG 22


RESULT  38
I49791/c
LOCUS       I49791          26 bp    DNA             PAT       07-OCT-1997
DEFINITION  Sequence 14 from patent US 5641661.
ACCESSION   I49791
VERSION     I49791.1  GI:2472011
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 26)
  AUTHORS   Kumagai,M.H. and Sverlow,G.G.
  TITLE     Pichia pastoris alcohol oxidase ZZA1 and ZZA2 regulatory regions
            for heterologous gene expression
  JOURNAL   Patent: US 5641661-A 14 24-JUN-1997;
FEATURES             Location/Qualifiers
     source          1. .26
                     /organism="unknown"
BASE COUNT         9 a      9 c      4 g      4 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 6;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      22 GACGTTCG 15


RESULT  39
AR121792/c
LOCUS       AR121792        27 bp    DNA             PAT       16-MAY-2001
DEFINITION  Sequence 27 from patent US 6160099.
ACCESSION   AR121792
VERSION     AR121792.1  GI:14105368
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 27)
  AUTHORS   Jonak,Z.Ludmila, Taylor,A., Trulli,S.H and Johanson,K.O.
  TITLE     Anti-human .alpha. .sub.v .beta. .sub.3 and .alpha. .sub.v
            .beta. .sub.5 antibodies
  JOURNAL   Patent: US 6160099-A 27 12-DEC-2000;
FEATURES             Location/Qualifiers
     source          1. .27
                     /organism="unknown"
BASE COUNT         7 a      10 c      7 g      3 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 6;  Length 27;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      22 GACGTTCG 15


RESULT  40
AR001148
LOCUS       AR001148        31 bp    DNA             PAT       04-DEC-1998
DEFINITION  Sequence 12 from patent US 5738990.
ACCESSION   AR001148
VERSION     AR001148.1  GI:3963215
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 31)
  AUTHORS   Edwards,C.A., Fry,K.E., Cantor,C.R. and Andrews,B.M.
  TITLE     Sequence-directed DNA-binding molecules compositions and methods
  JOURNAL   Patent: US 5738990-A 12 14-APR-1998;
FEATURES             Location/Qualifiers
     source          1. .31
                     /organism="unknown"
BASE COUNT         6 a      10 c      7 g      8 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 6;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      10 GACGTTCG 17


RESULT  41
AR003026
LOCUS       AR003026        31 bp    DNA             PAT       04-DEC-1998
DEFINITION  Sequence 12 from patent US 5744131.
ACCESSION   AR003026
VERSION     AR003026.1  GI:3964285
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 31)
  AUTHORS   Edwards,C.A., Fry,K.E., Cantor,C.R. and Andrews,B.M.
  TITLE     Sequence-directed DNA-binding molecules compositions and methods
  JOURNAL   Patent: US 5744131-A 12 28-APR-1998;
FEATURES             Location/Qualifiers
     source          1. .31
                     /organism="unknown"
BASE COUNT         6 a      10 c      7 g      8 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 6;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      10 GACGTTCG 17


RESULT  42
AR033000
LOCUS       AR033000        31 bp    DNA             PAT       29-SEP-1999
DEFINITION  Sequence 612 from patent US 5869241.
ACCESSION   AR033000
VERSION     AR033000.1  GI:5948605
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
```

```
REFERENCE   1  (bases 1 to 31)
  AUTHORS   Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
  TITLE     Method of determining DNA sequence preference of a DNA-binding
            molecule
  JOURNAL   Patent: US 5869241-A 612 09-FEB-1999;
FEATURES             Location/Qualifiers
     source          1. .31
                     /organism="unknown"
BASE COUNT        6 a      10 c       7 g       8 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 6;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      10 GACGTTCG 17


RESULT  43
AR126490
LOCUS       AR126490        31 bp    DNA             PAT       16-MAY-2001
DEFINITION  Sequence 117 from patent US 6180341.
ACCESSION   AR126490
VERSION     AR126490.1  GI:14113083
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 31)
  AUTHORS   Iverson,B.L., Georgiou,G. and Burks,E.A.
  TITLE     In vitro scanning saturation mutagenesis of proteins
  JOURNAL   Patent: US 6180341-A 117 30-JAN-2001;
FEATURES             Location/Qualifiers
     source          1. .31
                     /organism="unknown"
BASE COUNT        9 a      10 c       6 g       6 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 6;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      24 GACGTTCG 31


RESULT  44
AR126494
LOCUS       AR126494        31 bp    DNA             PAT       16-MAY-2001
DEFINITION  Sequence 121 from patent US 6180341.
ACCESSION   AR126494
VERSION     AR126494.1  GI:14113087
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 31)
  AUTHORS   Iverson,B.L., Georgiou,G. and Burks,E.A.
  TITLE     In vitro scanning saturation mutagenesis of proteins
  JOURNAL   Patent: US 6180341-A 121 30-JAN-2001;
FEATURES             Location/Qualifiers
     source          1. .31
                     /organism="unknown"
BASE COUNT        8 a      10 c       6 g       7 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 6;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      24 GACGTTCG 31


RESULT  45
I76870
LOCUS       I76870          31 bp    DNA             PAT       03-APR-1998
DEFINITION  Sequence 12 from patent US 5693463.
ACCESSION   I76870
VERSION     I76870.1  GI:3013024
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 31)
  AUTHORS   Edwards,C.A., Fry,K.E., Cantor,C.R. and Andrews,B.M.
  TITLE     Method of ordering sequence binding preferences of a DNA-binding
            molecule
  JOURNAL   Patent: US 5693463-A 12 02-DEC-1997;
FEATURES             Location/Qualifiers
     source          1. .31
                     /organism="unknown"
BASE COUNT        6 a      10 c       7 g       8 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 6;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      10 GACGTTCG 17


Search completed: November 29, 2001, 14:47:07
Job time: 8320 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: November 29, 2001, 14:51:04 ; Search time 158.03 Seconds
(without alignments)
43.401 Million cell updates/sec

Title: FRAG2
Perfect score: 8
Sequence: 1 GACGTTCG 8

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1084414

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_1101:*
1: /SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	8	100.0	14	19	AAV42287	Clone F4.1.8 kappa
c	2	8	100.0	17	15	AAQ58511	Sequence of primer
c	3	8	100.0	17	18	AAT88305	Oligonucleotide pr
c	4	8	100.0	17	20	AAX55861	PCR upstream prime
	5	8	100.0	17	22	AAF95099	Wild-type capture
	6	8	100.0	18	15	AAQ71408	Primer Df8 for hou
	7	8	100.0	18	22	AAH44096	Oryza sativa perox
c	8	8	100.0	19	17	AAT30550	Probe Jkappa1 for
	9	8	100.0	19	21	AAA86282	PCBA HH ribozyme b
	10	8	100.0	19	21	AAA86283	PCBA HH ribozyme b
	11	8	100.0	19	22	AAH61444	PCNA HH ribozyme b
	12	8	100.0	19	22	AAH61445	PCNA HH ribozyme b
c	13	8	100.0	20	17	AAT31844	Probe/Primer used
c	14	8	100.0	20	17	AAT28759	Probe #3 for Mycob
c	15	8	100.0	20	17	AAT06576	Probe B (Set 1) fo
c	16	8	100.0	20	17	AAT06500	Probe B (Set 1) fo
	17	8	100.0	20	18	AAT74311	Coprinus peroxidas
	18	8	100.0	20	19	AAV45999	Immune adjuvant CR
c	19	8	100.0	20	19	AAV31177	Bacillus thuringie
c	20	8	100.0	20	19	AAV21458	Mycobacterium tube
	21	8	100.0	20	20	AAZ06010	PCR primer used to
	22	8	100.0	20	20	AAX93946	PCR primer used to
	23	8	100.0	20	20	AAV80112	Oligo used in expe
c	24	8	100.0	20	22	AAH23258	Human MMIF mRNA in
	25	8	100.0	20	22	AAF99392	Immunostimulatory
	26	8	100.0	20	22	AAF99395	Immunostimulatory
	27	8	100.0	20	22	AAF99397	Immunostimulatory
	28	8	100.0	20	22	AAF99651	Immunostimulatory
	29	8	100.0	20	22	AAF99692	Immunostimulatory
c	30	8	100.0	20	22	AAF60653	HSV-1 R15 PCR prim
c	31	8	100.0	20	22	AAF60804	S. cerevisiae MET1
c	32	8	100.0	21	20	AAX34275	Primer AsdF2 for P
	33	8	100.0	22	22	AAF98750	Human IFN-alpha im
	34	8	100.0	22	22	AAF98873	Immunostimulatory
	35	8	100.0	22	22	AAF99680	Immunostimulatory
	36	8	100.0	22	22	AAF99776	Immunostimulatory
	37	8	100.0	22	22	AAF99832	Immunostimulatory
	38	8	100.0	23	17	AAT31845	Probe/Primer used
	39	8	100.0	23	17	AAT28760	Probe #4 for Mycob
c	40	8	100.0	23	22	AAD07148	Back module BM-379
c	41	8	100.0	23	22	AAD07155	Back module BM-654
c	42	8	100.0	23	22	AAF74942	Bacteriophage lamb
c	43	8	100.0	23	22	AAF74949	Bacteriophage lamb
	44	8	100.0	24	15	AAQ55886	Probe for OriS sit
	45	8	100.0	24	17	AAT06577	Probe B' (Set 1) f

ALIGNMENTS

RESULT 1
AAV42287
ID AAV42287 standard; cDNA; 14 BP.
XX
AC AAV42287;
XX
DT 23-SEP-1998 (first entry)
XX
DE Clone F4.1.8 kappa light chain transcript segemnt J-kappa.
XX
KW Human; immunoglobulin; Ig; transgenic; non-human mammal;
KW inactivated endogenous Ig locus; B-cell development;
KW human heavy chain Ig locus; micro constant region; J-H; D-H; V-H gene;
KW kappa light chain Ig locus; kappa constant region; J-kappa gene; V-kappa;
KW production; antibody; ss.
XX
OS Homo sapiens.
XX
PN WO9824893-A2.
XX
PD 11-JUN-1998.
XX
PF 03-DEC-1997; 97WO-US23091.
XX
PR 03-DEC-1996; 96US-0759620.
XX
PA (ABGE-) ABGENIX INC.
XX
PI Green L, Jakobovits A, Klapholz S, Kucherlapati R;
PI Mendez M;
XX
DR WPI; 1998-333314/29.
XX
PT New transgenic non-human mammals - having an inactivated

```
PT   immunoglobulin locus and a near complete human immunoglobulin locus,
PT   used for production of human antibodies
XX
PS   Example 8; Page 39; 128pp; English.
XX
CC   AAV42284-99 represent human kappa light chain immunoglobulin (Ig)
CC   transcripts expressed in XenoMouse II strains. The Xenomice were
CC   produced using the method of the invention. The specification describes
CC   a transgenic non-human mammal which has genome modifications that
CC   comprise an inactivated endogenous Ig locus, so that the mammal does not
CC   display normal B-cell development. The modified genome also has an
CC   inserted human heavy chain Ig locus in germline configuration, the human
CC   heavy chain Ig locus comprising a human micro constant region and
CC   regulatory and switch sequences, human J-H genes, human D-H genes, and
CC   human V-H genes and an inserted human kappa light chain Ig locus in
CC   germline configuration, the human kappa light chain Ig locus comprising a
CC   human kappa constant region, J-kappa genes, and V-kappa genes, where the
CC   number of V-H and V-kappa genes inserted are selected to restore normal
CC   B-cell development in the mammal. The transgenic animals have a near
CC   complete human Ig locus, including both a human heavy chain locus and a
CC   human kappa light chain locus. They can be used for the production of
CC   human antibodies when exposed to particular antigens e.g. when exposed to
CC   human IL-8, EGFR or TNF- alpha the mice will produce antibodies to IL-8,
CC   EGFR or TNF- alpha respectively.
XX
SQ   Sequence 14 BP; 3 A; 4 C; 5 G; 2 T; 0 other;


  Query Match             100.0%;  Score 8;  DB 19;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 8e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db       2 gacgttcg 9


RESULT   2
AAQ58511/c
ID   AAQ58511 standard; DNA; 17 BP.
XX
AC   AAQ58511;
XX
DT   12-SEP-1994  (first entry)
XX
DE   Sequence of primer HVRNA2F3 for the synthesis of P64 or P71
DE   variant - 5C - of Heliothis armigera RNA 2.
XX
KW   Heliothis armigera stunt virus; HaSV; RNA 2; small RNA virus;
KW   PCR primer; 5C version; ss.
XX
OS   Synthetic.
XX
PN   WO9404660-A.
XX
PD   03-MAR-1994.
XX
PF   13-AUG-1993;   93WO-AU00411.
XX
PR   14-AUG-1992;   92AU-0004081.
PR   08-JUL-1993;   93US-0089372.
XX
PA   (CSIR ) COMMONWEALTH SCI & IND RES ORG.
PA   (PACI-) PACIFIC SEEDS PTY LTD.
XX
PI   Christian PD,  Gordon KHJ,  Hanzlik TN;
XX
DR   WPI; 1994-083180/10.
XX
PT   Small RNA virus capable of infecting insect species, e.g.
PT   Heliothis - and transgenic plants contg. viral nucleic acid, for
PT   protection against insect pests
XX
PS   Claim 18; Page 130; 183pp; English.
XX
CC   H. armigera larvae were raised and viral RNA was extracted. The virus
CC   RNAs were reverse transcribed into cDNA. Clone E3 represents 99.7%
CC   or RNA 1. hr236 contains about 88% or RNA 2. hr236 was used as a
CC   basis for constructing the full length clone of RNA 2. The
CC   major ORFs encode the capsid protein precursor P71, and P17. P71
CC   is the precursor of P64. There is a variant of the major capsid
CC   protein - the 5C version - which carries an extra C residue. The
CC   carboxy terminal halves of the major capsid protein variant,
CC   whether terminating as for P64 or for P71, were produced using PCR.
CC   An oligo primer, HVRNA2F3, corresp. to bps 873-889 of AAQ58523, was
CC   used in conjunction with HVP65C and HVP6C2.
XX
SQ   Sequence 17 BP; 6 A; 3 C; 6 G; 2 T; 0 other;


  Query Match             100.0%;  Score 8;  DB 15;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 7.9e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      12 GACGTTCG 5


RESULT   3
AAT88305/c
ID   AAT88305 standard; DNA; 17 BP.
XX
AC   AAT88305;
XX
DT   22-JAN-1998  (first entry)
XX
DE   Oligonucleotide primer O_L3L_5.
XX
KW   Oligonucleotide primer; preparation; library; CDR3;
KW   complementarity determining region; ss.
XX
OS   Synthetic.
XX
PN   WO9708320-A1.
XX
PD   06-MAR-1997.
XX
PF   19-AUG-1996;   96WO-EP03647.
XX
PR   18-AUG-1995;   95EP-0113021.
XX
PA   (MORP-) MORPHOSYS GES PROTEINOPTIMIERUNG MBH.
XX
PI   Ge L,  Ilag V,  Knappik A,  Moroney S,  Pack P,  Plueckthun A;
XX
DR   WPI; 1997-179277/16.
XX
PT   Preparation of human derived antibody gene library - using synthetic
PT   consensus sequences, and signal consensus antibody gene as universal
PT   framework for highly diverse antibody libraries
XX
PS   Example 5; Page 39; 436pp; English.
XX
CC   The present sequence is an oligonucleotide primer used in the
CC   preparation of complementarity determining region 3 (CDR3)
CC   libraries.
XX
SQ   Sequence 17 BP; 5 A; 5 C; 6 G; 1 T; 0 other;


  Query Match             100.0%;  Score 8;  DB 18;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 7.9e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
```

```
          ||||||||
Db      16 GACGTTCG 9


RESULT   4
AAX55861/c
ID   AAX55861 standard; DNA; 17 BP.
XX
AC   AAX55861;
XX
DT   09-JUL-1999  (first entry)
XX
DE   PCR upstream primer #640 from WO9918240 Example 9.
XX
KW   Labelling; tag; molecular species; identification; property;
KW   characteristic; hybridisation; amplification; PCR primer; ss.
XX
OS   Synthetic.
XX
PN   WO9918240-A2.
XX
PD   15-APR-1999.
XX
PF   05-OCT-1998;   98WO-US20874.
XX
PR   06-OCT-1997;   97US-0944410.
XX
PA   (STRA-) STRATAGENE.
XX
PI   Sorge JA;
XX
DR   WPI; 1999-264040/22.
XX
PT   Uniquely tagged molecules identifiable by a unique property or
PT   characteristic
XX
PS   Example 9; Page 107; 138pp; English.
XX
CC   The present invention describes a composition comprising a mixture of
CC   different species of molecules where each species is linked to a tag
CC   that is unique to that species and that encodes at least two variable
CC   positions on that species, where the tags can be identified without the
CC   need for first isolating each of the tags prior to identification. Liquid
CC   phase hybridisation system may be used for simultaneous identification
CC   of a large subset of targets out of a very large collection of similar
CC   of dissimilar molecular species. It may also be used to create tagged
CC   molecules that identify any collection of molecular species, e.g.
CC   peptides, antibodies, nucleic acids. Method bar codes collections or
CC   probes or analytes for use in a liquid phase hybridisation method. Tagged
CC   probes able to detect small changes or mutations in the target specimen.
CC   Use of molecular tags overcomes difficulties of prior art methods, e.g.
CC   the concentration of the probe would not be limited by the solid support,
CC   both the target nucleic acids and the probes can diffuse toward each
CC   other, and signal amplification through cycling reactions could occur.
CC   Sequencing DNA with tags in combination with DNA amplification techniques
CC   means that there is no need for traditional sequencing methods or
CC   attaching to a solid phase, either the materials to be analysed or the
CC   tags. The present sequence represents a PCR primer which is used in an
CC   example from the present invention.
XX
SQ   Sequence 17 BP; 6 A; 5 C; 4 G; 2 T; 0 other;


  Query Match             100.0%;  Score 8;  DB 20;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 7.9e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      14 GACGTTCG 7


RESULT   5
AAF95099
ID   AAF95099 standard; DNA; 17 BP.
XX
AC   AAF95099;
XX
DT   23-MAY-2001  (first entry)
XX
DE   Wild-type capture oligonucleotide #26.
XX
KW   Tubercle bacillus; drug sensitivity; drug resistance; rifampicin;
KW   streptomycin; kanamycin; isoniazid; ethambutol; rpoB gene; rrs gene;
KW   rpsL gene; inhA gene; katG gene; embB gene; probe; PCR primer; ss.
XX
OS   Mycobacterium tuberculosis.
XX
PN   EP1076099-A2.
XX
PD   14-FEB-2001.
XX
PF   02-AUG-2000; 2000EP-0306563.
XX
PR   03-AUG-1999;   99JP-0220357.
XX
PA   (NISN ) NISSHINBO IND INC.
PA   (SYST-) SYSTEM RES INC.
XX
PI   Suzuki Y,  Nishida M,  Takenishi S;
XX
DR   WPI; 2001-246696/26.
XX
PT   New oligonucleotides, nucleic acid probes and primers are useful for
PT   differentiating drug-resistance and determining infection with tubercle
PT   bacilli -
XX
PS   Claim 25; Page 44; 114pp; English.
XX
CC   The present invention relates to oligonucleotides based on nucleotide
CC   sequences obtained from both wild-type tubercle bacilli (wtTB) that are
CC   susceptible to a drug and mutant-type tubercle bacilli (mtTB) that are
CC   resistant to a drug. The drugs used in the present invention are
CC   rifampicin (RFP), streptomycin (SM), kanamycin (KM), isoniazid (INH) and
CC   ethambutol (EB). The rpoB gene is responsible for resistance to RFP; the
CC   rrs gene is responsible for resistance to SM and KM; the rpsL gene is
CC   responsible for resistance to SM; the inhA gene is responsible for
CC   resistance to INH; the katG gene is responsible for resistance to INH;
CC   and the embB gene is responsible for resistance to EB. The present
CC   invention also relates to nucleic acid probes having part of a nucleotide
CC   sequence of tubercle bacilli (TB) responsible for drug resistance and
CC   primers used to generate the probes. The present sequence is an
CC   oligonucleotide of the present invention. The oligonucleotides of the
CC   present invention can be used to enable the differentiation of drug
CC   resistance and the determination of infection with tubercle bacilli
CC   simultaneously.
XX
SQ   Sequence 17 BP; 3 A; 4 C; 5 G; 5 T; 0 other;


  Query Match             100.0%;  Score 8;  DB 22;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 7.9e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db       7 gacgttcg 14


RESULT   6
AAQ71408
ID   AAQ71408 standard; cDNA; 18 BP.
XX
AC   AAQ71408;
XX
DT   01-APR-1995  (first entry)
```

XX
DE Primer Df8 for house dust mite allergen DerfVII cDNA.
XX
KW Primer; Df8; DNA sequencing; DerfVII allergen; antiallergic;
KW allergy diagnosis; ss.
XX
OS Dermatophagoides farinae.
XX
PN WO9420614-A.
XX
PD 15-SEP-1994.
XX
PF 11-MAR-1994; 94WO-AU00117.
XX
PR 12-MAR-1993; 93US-0031141.
PR 22-JUN-1993; 93US-0081540.
XX
PA (CHIL-) INST CHILD HEALTH RES.
XX
PI Chua K, Thomas WR;
XX
DR WPI; 1994-303021/37.
XX
PT New nucleic acid encoding specific dust mite allergens - and
PT related vectors, transformed cells, peptides and antibodies,
PT useful for desensitisation and diagnosis.
XX
PS Example 5; Page 33; 67pp; English.
XX
CC The DNA sequencing primer Df8 is derived from nucleotides 225-208
CC of DerfVII (AAQ71401) and is used in the polymerase chain reaction
CC amplification and sequencing of a cDNA clone encoding DerfVII from a
CC phage lambda-gt11 cDNA library. DerfVII antigen is useful as an
CC antiallergic reagent for treating sensitivity to house dust mite
CC allergens.
XX
SQ Sequence 18 BP; 5 A; 3 C; 3 G; 7 T; 0 other;

Query Match 100.0%; Score 8; DB 15; Length 18;
Best Local Similarity 100.0%; Pred. No. 7.8e+03;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGTTCG 8
||||||||
Db 3 gacgttcg 10

RESULT 7
AAH44096
ID AAH44096 standard; DNA; 18 BP.
XX
AC AAH44096;
XX
DT 12-SEP-2001 (first entry)
XX
DE Oryza sativa peroxidase PCR primer/probe SEQ ID NO:47.
XX
KW Oryza sativa; rice; peroxidase; POX; characteristic; gene expression;
KW modification; plant; bacterial infection; Magnaporthe grisea; probe;
KW PCR primer; ss.
XX
OS Oryza sativa.
XX
PN WO200142475-A1.
XX
PD 14-JUN-2001.
XX
PF 08-DEC-2000; 2000WO-JP08728.
XX
PR 10-DEC-1999; 99JP-0352472.
XX
PA (NORQ) JAPAN MIN AGRIC FORESTRY & FISHERIES.
XX
PI Ohashi Y, Mitsuhara I, Sasaki T, Nagamura Y, Ito H, Iwai T;
PI Hiraga S;
XX
DR WPI; 2001-381695/40.
XX
PT New set of rice peroxidase genes for analysis of peroxidase expression
PT in rice under varying conditions and production of rice plants with
PT desired characteristics -
XX
PS Example 2; Page 62; 258pp; Japanese.
XX
CC The present invention describes a set of peroxidase genes found in
CC plants, especially rice, and their homologues, modified forms and
CC fragments, where the sequences of the peroxidase genes in the set are
CC given in AAH44071 to AAH44091. Also described are: (1) promoters for the
CC control of the gene set; (2) the preparation of cassette vectors using
CC the genes and promoters; (3) analysis of plant characteristics using the
CC peroxidase set by isolating RNA from the plant, binding the RNA to a
CC membrane, mixing with a labelled peroxidase gene set, incubating, and
CC detecting the label signal to show which genes in the set are expressed
CC in the sample plant; and (4) DNA microarrays for peroxidase gene
CC expression analysis. The set of genes are used for the analysis of the
CC pattern of peroxidase gene expression in particular rice plants and
CC their component tissues and under different environmental conditions,
CC and modification of rice plants to provide desired specificities of
CC peroxidase gene expression to impart particular characteristics to the
CC plants such as response to bacterial infection by Magnaporthe grisea.
CC The present sequence represents a PCR primer/probe for a rice peroxidase
CC (POX) gene, which is used in an example from the present invention.
XX
SQ Sequence 18 BP; 4 A; 5 C; 5 G; 4 T; 0 other;

Query Match 100.0%; Score 8; DB 22; Length 18;
Best Local Similarity 100.0%; Pred. No. 7.8e+03;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGTTCG 8
||||||||
Db 5 gacgttcg 12

RESULT 8
AAT30550/c
ID AAT30550 standard; DNA; 19 BP.
XX
AC AAT30550;
XX
DT 11-FEB-1997 (first entry)
XX
DE Probe Jkappa1 for HNK-20 Jkappa chain coding sequence.
XX
KW Antibody; HNK-20; variable heavy chain; hybridoma; murine; IgA; mouse;
KW F glycoprotein; respiratory syncytial virus; RSV; constant region gene;
KW chimeric antibody; isotype-switched antibody; therapy; infection; human;
KW pneumonia; bronchiolitis; animal; polymerase chain reaction; probe; ss.
XX
OS Synthetic.
XX
PN WO9616974-A1.
XX
PD 06-JUN-1996.
XX
PF 01-DEC-1995; 95WO-US15716.
XX
PR 01-DEC-1994; 94US-0348548.
XX
PA (ORAV-) ORAVAX INC.
XX
PI Berdoz J, Kraehenbuhl J;
XX
DR WPI; 1996-286826/29.

```
XX
PT   DNA encoding variable region of antibody HNK-20 - for treating
PT   respiratory syncytial virus infection
XX
PS   Example; Page 53; 75pp; English.
XX
CC   AAT30546-T30558 represent probes for the J chains of an antibody
CC   produced by the hybridoma cell HNK-20.  AAT30550-T30554 represent probes
CC   for the Jkappa chain of the HNK-20 antibody.  HNK-20 is a murine
CC   hybridoma cell line, that produces IgA specific for the F glycoprotein of
CC   respiratory syncytial virus (RSV).  The variable chain coding sequences
CC   (see AAT30456-T30458) were isolated using primers specific for the 5'
CC   untranslated region of the variable region, and for the intron
CC   downstream of the rearranged J region (see AAT30459-T30545).  The
CC   amplified sequences can be inserted into vectors containing heterologous
CC   (such as human) constant region genes, for the production of chimeric
CC   and isotype-switched antibodies.  The antibodies are useful in the
CC   treatment and diagnosis of infection by RSV, such as pneumonia and
CC   bronchiolitis, in humans and animals.  By using genomic DNA as a
CC   template, variable region genes can be isolated without producing
CC   fragments that have to be adapted for recombinant antibody expression.
CC   Also, by using the genomic DNA, no knowledge of the DNA sequence encoding
CC   the target variable region is required.  Chimeric antibodies produced
CC   from the encoded proteins, that contain the constant region of the host
CC   being treated, are less likely to cause adverse immune reactions.
XX
SQ   Sequence 19 BP; 3 A; 9 C; 4 G; 3 T; 0 other;

  Query Match             100.0%;  Score 8;  DB 17;  Length 19;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      18 GACGTTCG 11


RESULT   9
AAA86282
ID   AAA86282 standard; DNA; 19 BP.
XX
AC   AAA86282;
XX
DT   04-DEC-2000  (first entry)
XX
DE   PCBA HH ribozyme binding site #14.
XX
KW   Ribozyme; hairpin; hammerhead; gene therapy; vasotropic;
KW   restenosis; ss.
XX
OS   Mammalia.
XX
PN   WO200032765-A2.
XX
PD   08-JUN-2000.
XX
PF   06-DEC-1999;   99WO-US28772.
XX
PR   04-DEC-1998;   98US-0110954.
XX
PA   (IMMU-) IMMUSOL INC.
XX
PI   Tritz R,  Welch PJ,  Barber JR,  Robbins JM;
XX
DR   WPI; 2000-412314/35.
XX
PT   New hairpin and hammerhead ribozyme for inhibiting restenosis, cleaves
PT   RNA encoding a cyclin or cell-cycle dependent kinase other than CDK1,
PT   PCNA and Cyclin B1  -
XX
PS   Disclosure; Page 105; 109pp; English.
XX
CC   The present invention relates to a hairpin or hammerhead ribozyme,
CC   designed to cleave RNA encoding a cyclin or cell-cycle dependent kinase
CC   other than cell-cycle dependent kinases CDK1, PCNA and Cyclin B1.
CC   Representative examples of ribozyme recognition sites are given in
CC   AAA82415 to AAA86787. The ribozyme of the invention is useful for
CC   inhibiting restenosis by introduction of the ribozyme into cells.
CC   The ribozyme is resistant to endonuclease activity and hence is
CC   efficient in restenosis treatment.
XX
SQ   Sequence 19 BP; 1 A; 9 C; 5 G; 4 T; 0 other;

  Query Match             100.0%;  Score 8;  DB 21;  Length 19;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db       5 gacgttcg 12


RESULT  10
AAA86283
ID   AAA86283 standard; DNA; 19 BP.
XX
AC   AAA86283;
XX
DT   04-DEC-2000  (first entry)
XX
DE   PCBA HH ribozyme binding site #15.
XX
KW   Ribozyme; hairpin; hammerhead; gene therapy; vasotropic;
KW   restenosis; ss.
XX
OS   Mammalia.
XX
PN   WO200032765-A2.
XX
PD   08-JUN-2000.
XX
PF   06-DEC-1999;   99WO-US28772.
XX
PR   04-DEC-1998;   98US-0110954.
XX
PA   (IMMU-) IMMUSOL INC.
XX
PI   Tritz R,  Welch PJ,  Barber JR,  Robbins JM;
XX
DR   WPI; 2000-412314/35.
XX
PT   New hairpin and hammerhead ribozyme for inhibiting restenosis, cleaves
PT   RNA encoding a cyclin or cell-cycle dependent kinase other than CDK1,
PT   PCNA and Cyclin B1  -
XX
PS   Disclosure; Page 105; 109pp; English.
XX
CC   The present invention relates to a hairpin or hammerhead ribozyme,
CC   designed to cleave RNA encoding a cyclin or cell-cycle dependent kinase
CC   other than cell-cycle dependent kinases CDK1, PCNA and Cyclin B1.
CC   Representative examples of ribozyme recognition sites are given in
CC   AAA82415 to AAA86787. The ribozyme of the invention is useful for
CC   inhibiting restenosis by introduction of the ribozyme into cells.
CC   The ribozyme is resistant to endonuclease activity and hence is
CC   efficient in restenosis treatment.
XX
SQ   Sequence 19 BP; 1 A; 9 C; 5 G; 4 T; 0 other;

  Query Match             100.0%;  Score 8;  DB 21;  Length 19;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
```

```
          ||||||||
Db        4 gacgttcg 11


RESULT  11
AAH61444
ID   AAH61444 standard; DNA; 19 BP.
XX
AC   AAH61444;
XX
DT   10-SEP-2001  (first entry)
XX
DE   PCNA HH ribozyme binding site SEQ ID NO:3868.
XX
KW   Human; ribozyme therapy; hairpin ribozyme; hammerhead ribozyme;
KW   recognition site; target; ribozyme binding site; eye disease; vulnerary;
KW   proliferative disease; skin disease; psoriasis; diabetic retinopathy;
KW   cytokine; inflammation; cell-cycle dependent kinase; cyclin; MMP;
KW   matrix metalloproteinase; growth factor; reductase; scarring; cytostatic;
KW   antipsoriatic; dermatological; antiseborrheic; antidiabetic; virucide;
KW   antisickling; ophthalmological; keratolytic; gene therapy; viral wart;
KW   atopic dermatitis; actinic keratosis; squamous cell carcinoma;
KW   basal cell carcinoma; seborrheic wart; vitreoretinopathy; scar;
KW   sickle cell retinopathy; ss.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200130362-A2.
XX
PD   03-MAY-2001.
XX
PF   26-OCT-2000; 2000WO-US29500.
XX
PR   26-OCT-1999;   99US-0161532.
XX
PA   (IMMU-) IMMUSOL INC.
XX
PI   Robbins JM,  Tritz R;
XX
DR   WPI; 2001-300427/31.
XX
PT   Treating proliferative skin or eye diseases and scarring, using
PT   ribozymes that cleave RNA encoding cytokines involved in inflammation,
PT   matrix metalloproteinases, growth factors and cell-cycle dependent
PT   kinases -
XX
PS   Example 1; Page 353; 408pp; English.
XX
CC   The present invention describes a method for treating a proliferative
CC   skin or eye disease and scarring. The method involves administering a
CC   ribozyme (I) which cleaves RNA encoding a cytokine involved in
CC   inflammation, matrix metalloproteinase (MMP), cyclin, cell-cycle
CC   dependent kinase, growth factor or a reductase, or administering a
CC   nucleic acid molecule (II) comprising a promoter operably linked to a
CC   nucleic acid segment encoding (I). (I) can have antipsoriatic,
CC   dermatological, cytostatic, antiseborrheic, antidiabetic, antisickling,
CC   ophthalmological, vulnerary, keratolytic and virucide activities, and
CC   cleaves RNA encoding cytokine involved in inflammation. (I) can be used
CC   in gene therapy. (I) and (II) are useful for treating proliferative
CC   skin diseases such as psoriasis, atopic dermatitis, actinic keratosis,
CC   squamous or basal cell carcinoma and viral or seborrheic wart. They can
CC   also be used for treating proliferative eye diseases such as diabetic
CC   retinopathy, vitreoretinopathy, sickle cell retinopathy, retinopathy of
CC   prematurity and retinal detachment, and for treating and preventing
CC   scarring such as keloid, adhesion and hypertrophic or hypertrophic burn
CC   scar. AAH57577 to AAH62099 represent sequences used in the
CC   exemplification of the present invention.
XX
SQ   Sequence 19 BP; 1 A; 9 C; 5 G; 4 T; 0 other;

Query Match             100.0%;  Score 8;  DB 22;  Length 19;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 GACGTTCG 8
            ||||||||
Db        5 gacgttcg 12


RESULT  12
AAH61445
ID   AAH61445 standard; DNA; 19 BP.
XX
AC   AAH61445;
XX
DT   10-SEP-2001  (first entry)
XX
DE   PCNA HH ribozyme binding site SEQ ID NO:3869.
XX
KW   Human; ribozyme therapy; hairpin ribozyme; hammerhead ribozyme;
KW   recognition site; target; ribozyme binding site; eye disease; vulnerary;
KW   proliferative disease; skin disease; psoriasis; diabetic retinopathy;
KW   cytokine; inflammation; cell-cycle dependent kinase; cyclin; MMP;
KW   matrix metalloproteinase; growth factor; reductase; scarring; cytostatic;
KW   antipsoriatic; dermatological; antiseborrheic; antidiabetic; virucide;
KW   antisickling; ophthalmological; keratolytic; gene therapy; viral wart;
KW   atopic dermatitis; actinic keratosis; squamous cell carcinoma;
KW   basal cell carcinoma; seborrheic wart; vitreoretinopathy; scar;
KW   sickle cell retinopathy; ss.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200130362-A2.
XX
PD   03-MAY-2001.
XX
PF   26-OCT-2000; 2000WO-US29500.
XX
PR   26-OCT-1999;   99US-0161532.
XX
PA   (IMMU-) IMMUSOL INC.
XX
PI   Robbins JM,  Tritz R;
XX
DR   WPI; 2001-300427/31.
XX
PT   Treating proliferative skin or eye diseases and scarring, using
PT   ribozymes that cleave RNA encoding cytokines involved in inflammation,
PT   matrix metalloproteinases, growth factors and cell-cycle dependent
PT   kinases -
XX
PS   Example 1; Page 353; 408pp; English.
XX
CC   The present invention describes a method for treating a proliferative
CC   skin or eye disease and scarring. The method involves administering a
CC   ribozyme (I) which cleaves RNA encoding a cytokine involved in
CC   inflammation, matrix metalloproteinase (MMP), cyclin, cell-cycle
CC   dependent kinase, growth factor or a reductase, or administering a
CC   nucleic acid molecule (II) comprising a promoter operably linked to a
CC   nucleic acid segment encoding (I). (I) can have antipsoriatic,
CC   dermatological, cytostatic, antiseborrheic, antidiabetic, antisickling,
CC   ophthalmological, vulnerary, keratolytic and virucide activities, and
CC   cleaves RNA encoding cytokine involved in inflammation. (I) can be used
CC   in gene therapy. (I) and (II) are useful for treating proliferative
CC   skin diseases such as psoriasis, atopic dermatitis, actinic keratosis,
CC   squamous or basal cell carcinoma and viral or seborrheic wart. They can
CC   also be used for treating proliferative eye diseases such as diabetic
CC   retinopathy, vitreoretinopathy, sickle cell retinopathy, retinopathy of
CC   prematurity and retinal detachment, and for treating and preventing
CC   scarring such as keloid, adhesion and hypertrophic or hypertrophic burn
CC   scar. AAH57577 to AAH62099 represent sequences used in the
CC   exemplification of the present invention.
XX
```

```
SQ   Sequence 19 BP; 1 A; 9 C; 5 G; 4 T; 0 other;

  Query Match             100.0%;  Score 8;  DB 22;  Length 19;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db       4 gacgttcg 11


RESULT  13
AAT31844/c
ID   AAT31844 standard; DNA; 20 BP.
XX
AC   AAT31844;
XX
DT   10-JAN-1997  (first entry)
XX
DE   Probe/Primer used for amplifying M. tuberculosis target sequence.
XX
KW   Sonication; probe; primer; Mycobacterium tuberculosis; analysis;
KW   detection; ss.
XX
OS   Synthetic.
XX
PN   WO9619301-A1.
XX
PD   27-JUN-1996.
XX
PF   22-DEC-1995;   95WO-US16810.
XX
PR   21-DEC-1995;   95US-0564995.
PR   22-DEC-1994;   94US-0362640.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Halaka FG;
XX
DR   WPI; 1996-309363/31.
XX
PT   Device for sonicating samples, e.g. body fluid sample - having
PT   electrical wave generator, vibrating element and vibratable member
PT   to generate standing sonic wave
XX
PS   Example 3; Page 26; 35pp; English.
XX
CC   A new device for sonicating a sample comprises an electrical wave
CC   generator, a vibrating element electrically connected to the
CC   electrical wave generator and a vibratable member transcersely
CC   secured to the vibrating element. The device can be used to lyse
CC   cells such as bacteria and viruses so that the intracellular
CC   components can be analysed. Four probes/primers (AAT31842-45) were
CC   used to amplify and detect a target sequence (AAT42980) generated from
CC   lysed Mycobacterium tuberculosis cells. These sequences were
CC   published to illustrate the invention.
XX
SQ   Sequence 20 BP; 6 A; 5 C; 6 G; 3 T; 0 other;

  Query Match             100.0%;  Score 8;  DB 17;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      18 GACGTTCG 11


RESULT  14
AAT28759/c
ID   AAT28759 standard; DNA; 20 BP.
XX
AC   AAT28759;
XX
DT   11-DEC-1996  (first entry)
XX
DE   Probe #3 for Mycobacterium tuberculosis protein antigen b gene.
XX
KW   Probe; Chlamydia trachomatis; amplification inhibition; cryptic plasmid;
KW   ligase chain reaction; LCR; primer; amplify; Mycobacterium tuberculosis;
KW   protein antigen b; pab; Opa A; Neisseria gonorrhoeae; spermidine;
KW   magnesium chloride; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1
FT                   /*tag= a
FT                   /note= "conjugated with adamantaneacetic acid
FT                          ("adamantane")"
XX
PN   WO9612824-A2.
XX
PD   02-MAY-1996.
XX
PF   18-OCT-1995;   95WO-US12874.
XX
PR   12-OCT-1995;   95US-0532212.
PR   21-OCT-1994;   94US-0331391.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Davis AH,  Lee EH;
XX
DR   WPI; 1996-230622/23.
XX
PT   Use of spermidine to relieve inhibition of ligase chain reaction -
PT   improves amplification yield and allows reaction to proceed in
PT   presence of non-optimal magnesium chloride concentrations
XX
PS   Example 1; Page 15; 33pp; English.
XX
CC   AAT28753-T28768 represent probes used in the method of the invention, to
CC   relieve amplification inhibition in a ligase chain reaction (LCR).
CC   AAT28757-T28760 represent probes for the protein antigen b (pab) gene of
CC   Mycobacterium tuberculosis. In the method of the invention, the LCR
CC   amplification inhibition is relieved by forming a reaction mixture
CC   containing a clinical test sample, an amount of spermidine effective to
CC   relieve amplification inhibition, and a composition containing two pairs
CC   of probes. Each pair of probes used contains a primary probe that is
CC   hybridisable to the target sequence. The two primary probes used are
CC   hybridisable to adjacent or near adjacent positions on the target
CC   sequence. The primary probes are then hybridised to the target, and
CC   ligated to form a fused product. The fused product is then dissociated
CC   from the target. One or more of the primary or secondary probes used
CC   can be modified at one end, so that it is non-ligatable to the other
CC   primary or secondary probe. The presence of spermidine alleviates the
CC   inhibitory effects of certain clinical samples on LCR reactions, and
CC   therefore improves amplification yields. The presence of spermidine also
CC   allows LCR to proceed in the presence of non-optimal concentrations of
CC   magnesium chloride. This means that LCR can be carried out in subsequent
CC   assay mixtures that require different concentrations of magnesium
CC   chloride.
XX
SQ   Sequence 20 BP; 6 A; 5 C; 6 G; 3 T; 0 other;

  Query Match             100.0%;  Score 8;  DB 17;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      18 GACGTTCG 11
```

```
RESULT 15
AAT06576/c
ID   AAT06576 standard; DNA; 20 BP.
XX
AC   AAT06576;
XX
DT   25-JUN-1996  (first entry)
XX
DE   Probe B (Set 1) for M. tuberculosis Pab gene nucleotides 347-390.
XX
KW   probe; modified ligase chain reaction; Mycobacterium tuberculosis;
KW   M. avium; M. intracellulare; M. kansasii; detection; diagnosis; ss.
XX
OS   Synthetic.
XX
PN   WO9531571-A2.
XX
PD   23-NOV-1995.
XX
PF   04-MAY-1995;   95WO-US05816.
XX
PR   13-MAY-1994;   94US-0223330.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Kratochvil JD,  Leckie GW,  Odonnell DL,  Solomon NA;
XX
DR   WPI; 1996-010956/01.
XX
PT   New probes for detection of Mycobacterium species - derived from the
PT   16S ribosomal RNA gene, the protein antigen b gene and the 65 kD and
PT   10 kD heat shock protein genes of M.tuberculosis
XX
PS   Example 1; Page 34; 60pp; English.
XX
CC   Probe set 1 (AAT06574-577) was selected to detect a target sequence in
CC   Mycobacterium tuberculosis corresponding to nucleotides 347-390
CC   (AAT06573) of the protein antigen b (pab) gene. The probes were labelled
CC   with carbazole and adamantane. Set 1 was capable of detecting as few as
CC   10 mols. of DNA derived from M. tuberculosis and showed no
CC   cross-reactivity with DNA genomes derived from M. avium, M.
CC   intracellulare and M.kansasii. A modified ligase chain reaction was
CC   utilised which uses two pairs of probes designated A, B (primary probes)
CC   and A', B' (secondary probes). Probe pairs were directed to the same
CC   target strand and ultimately ligated to one another after annealing to
CC   the target strand. At least one of the probes of a pair had a modified
CC   end with respect to the point of ligation.  The modified end had bases
CC   omitted to create a gap between one probe terminus and the next probe
CC   terminus when the pair was annealed to the target sequence.  Other
CC   modified ends include a base mismatched with the target sequence.  The
CC   presence of modified ends reduced the falsely positive signal created by
CC   blunt-end ligation of the complementary probe duplexes to one another in
CC   the absence of target. "Correction" of the modification, in a target
CC   dependent manner, was subsequently carried out to render the probes
CC   ligatable. Once ligated, the fused (reorganised) probe was dissociated
CC   (e.g. melted) from the target and, as with conventional LCR, the process
CC   was repeated for several cycles.
XX
SQ   Sequence 20 BP; 6 A; 5 C; 6 G; 3 T; 0 other;

  Query Match             100.0%;  Score 8;  DB 17;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
            ||||||||
Db      18 GACGTTCG 11


RESULT 16
```

```
AAT06500/c
ID   AAT06500 standard; DNA; 20 BP.
XX
AC   AAT06500;
XX
DT   25-JUN-1996  (first entry)
XX
DE   Probe B (Set 1) for M. tuberculosis Pab gene nucleotides 347-390.
XX
KW   probe; modified ligase chain reaction; Mycobacterium tuberculosis;
KW   M. avium; M. intracellulare; M. kansasii; detection; diagnosis; ss.
XX
OS   Synthetic.
XX
PN   WO9531570-A1.
XX
PD   23-NOV-1995.
XX
PF   04-MAY-1995;   95WO-US05602.
XX
PR   13-MAY-1994;   94US-0242403.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Davis AH,  Leckie GW,  Manlove MT,  Semple-Facey IE,  Solomon NA;
XX
DR   WPI; 1996-010955/01.
XX
PT   New probes for detection of M.tuberculosis - derived from e.g. the
PT   gene coding for protein antigen b and from the insertion-like
PT   element IS6110 of M.tuberculosis.
XX
PS   Claim 2; Page 34; 60pp; English.
XX
CC   Probe set 1 (AAT06498-501) was selected to detect a target sequence in
CC   Mycobacterium tuberculosis corresponding to nucleotides 347-390
CC   (AAT06502) of the protein antigen b (pab) gene. The probes were labelled
CC   with carbazole and adamantane. Set 1 was capable of detecting as few as
CC   10 mols. of DNA derived from M. tuberculosis and showed no
CC   cross-reactivity with DNA genomes derived from M. avium, M.
CC   intracellulare and M.kansasii. A modified ligase chain reaction was
CC   utilised which uses two pairs of probes designated A, B (primary probes)
CC   and A', B' (secondary probes). Probe pairs were directed to the same
CC   target strand and ultimately ligated to one another after annealing to
CC   the target strand. At least one of the probes of a pair had a modified
CC   end with respect to the point of ligation.  The modified end had bases
CC   omitted to create a gap between one probe terminus and the next probe
CC   terminus when the pair was annealed to the target sequence.  Other
CC   modified ends include a base mismatched with the target sequence.  The
CC   presence of modified ends reduced the falsely positive signal created by
CC   blunt-end ligation of the complementary probe duplexes to one another in
CC   the absence of target. "Correction" of the modification, in a target
CC   dependent manner, was subsequently carried out to render the probes
CC   ligatable. Once ligated, the fused (reorganised) probe was dissociated
CC   (e.g. melted) from the target and, as with conventional LCR, the process
CC   was repeated for several cycles.
XX
SQ   Sequence 20 BP; 6 A; 5 C; 6 G; 3 T; 0 other;

  Query Match             100.0%;  Score 8;  DB 17;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
            ||||||||
Db      18 GACGTTCG 11


RESULT 17
AAT74311
ID   AAT74311 standard; DNA; 20 BP.
XX
```

```
AC   AAT74311;
XX
DT   09-FEB-1998  (first entry)
XX
DE   Coprinus peroxidase PCR primer JC24.1.
XX
KW   Cleaning; bleaching; cellulose; fabric; enzyme hybrid; peroxidase;
KW   cellulose binding domain; Humicola insolens; cellulase;
KW   Myceliophthora thermophila; laccase; plasmid pJC24; PCR; primer;
KW   ss.
XX
OS   Synthetic.
OS   Coprinus cinereus.
XX
PN   WO9728243-A1.
XX
PD   07-AUG-1997.
XX
PF   29-JAN-1997;   97WO-DK00042.
XX
PR   29-JAN-1996;   96DK-0000094.
XX
PA   (NOVO ) NOVO-NORDISK AS.
XX
PI   Bjornvad ME,  Cherry JR,  Rasmussen MD,  Vind J,  Von Der Osten C;
XX
DR   WPI; 1997-402598/37.
XX
PT   Cleaning of cellulosic fabrics - using an enzyme hybrid comprising a
PT   sequence of a non-cellulolytic enzyme linked to a cellulose-binding
PT   domain sequence
XX
PS   Example 6; Page 84; 124pp; English.
XX
CC   PCR primers CiPpcrdwn (AAT74310) and JC24.1 (AAT74311) were used to
CC   amplify a DNA fragment containing the Coprinus cinereus peroxidase
CC   (CiP) signal sequence (22 amino acids), the Humicola insolens
CC   family 45 cellulase cellulose binding domain (CBD, 37 amino acids)
CC   and a NifA linker domain from Klebsiella pneumoniae (14 amino
CC   acids) using plasmid pJC23 (see AAT74280) as template. The PCR
CC   product was ligated to amplified laccase cDNA from Myceliophthora
CC   thermophila in vector pJC106 to obtain plasmid pJC24 (see AAT74281).
CC   Primer CiPpcrdwn was also used in the construction of plasmid pJC25
CC   (see AAT74282). A claimed process for removal or bleaching of soiling
CC   or stains on a cellulosic fabric comprises contacting the fabric
CC   with a modified enzyme (enzyme hybrid) comprising a catalytically
CC   active portion of a non-cellulolytic enzyme linked to a CBD. The
CC   hybrid enzyme gives improved enzyme performance by increasing the
CC   affinity of the enzyme for the fabric.
XX
SQ   Sequence 20 BP; 3 A; 5 C; 8 G; 4 T; 0 other;

  Query Match             100.0%;  Score 8;  DB 18;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db       9 gacgttcg 16


RESULT  18
AAV45999
ID   AAV45999 standard; DNA; 20 BP.
XX
AC   AAV45999;
XX
DT   16-OCT-1998  (first entry)
XX
DE   Immune adjuvant CRE-TC.
XX
KW   Immune system; adjuvant; vaccine; cancer; prophylactic; pathogenicity;
KW   modulator; tolerance; regulator; helper cell; antigen; immunoglobulin;
KW   Ig class; autoimmune response; T-cell; B-cell; tumour; ss.
XX
OS   Class Bacteria.
XX
PN   EP855184-A1.
XX
PD   29-JUL-1998.
XX
PF   23-JAN-1997;   97EP-0101019.
XX
PR   23-JAN-1997;   97EP-0101019.
XX
PA   (HEEG/) HEEG K.
PA   (LIPF/) LIPFORD G B.
PA   (WAGN/) WAGNER H.
XX
PI   Heeg K,  Lipford GB,  Wagner H;
XX
DR   WPI; 1998-389630/34.
XX
PT   Antigenic composition comprises polynucleotide fragment and antigen
PT   - used as vaccine to treat or prevent e.g. cancer or pathogen
PT   infections and to modulate immune response e.g. tolerance break and
PT   regulation of TH1/TH2 cells
XX
PS   Example 3; Page 7; 28pp; English.
XX
CC   AAV45993-V46019 are fragments of bacterial polynucleotides which are
CC   used as immune adjuvants for inclusion into vaccines to treat cancer and
CC   for prophylaxis and/or treatment of conditions caused by pathogenic
CC   micro-organisms. The polynucleotide is used for modulation of an immune
CC   response and the modulation is selected from the group break of
CC   tolerance, regulation of TH1/TH2 helper cell responses, switch of Ig
CC   classes, treatment of autoimmune responses and induction of tolerances.
CC   DNA oligomers are used to enhance the reactivity of immune cells to
CC   viral, bacterial and parasitic antigens, to break tolerance in anergic T
CC   and B cells e.g. against tumour antigens, as adjuvants in vaccination
CC   against tumour-defined antigens and immunostimulatory substances in an
CC   immune response against tumours and to suppress immune reactions of the
CC   innate and acquired immune system. The composition is inexpensive and
CC   stable and does not cause lethal shock, which happens with prior art
CC   bacterial sequences.
XX
SQ   Sequence 20 BP; 4 A; 5 C; 6 G; 5 T; 0 other;

  Query Match             100.0%;  Score 8;  DB 19;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db       8 gacgttcg 15


RESULT  19
AAV31177/c
ID   AAV31177 standard; DNA; 20 BP.
XX
AC   AAV31177;
XX
DT   28-SEP-1998  (first entry)
XX
DE   Bacillus thuringiensis Cry1 hybrid delta-endotoxin DNA exchange site.
XX
KW   hybrid; delta-endotoxin; Cry1; crystal protein; insecticide;
KW   plant protection; transgenic plants; resistance; DNA exchange site; ss.
XX
OS   Synthetic.
OS   Bacillus thuringiensis.
XX
PN   WO9822595-A1.
```

XX
PD 28-MAY-1998.
XX
PF 20-NOV-1997; 97WO-US21587.
XX
PR 03-SEP-1997; 97US-0922505.
PR 20-NOV-1996; 96US-0754490.
XX
PA (ECOG-) ECOGEN INC.
XX
PI Gilmer AJ, Malvar T;
XX
DR WPI; 1998-312480/27.
XX
PT New nucleic acid encoding Bacillus thuringiensis hybrid delta toxins
PT - with increased and broader spectrum activity, used to produce
PT transgenic plants with increased resistance to insects
XX
PS Example 1; Page 84; 209pp; English.
XX
CC The sequence is that of a DNA exchange site which was used in the
CC construction of a Bacillus thuringiensis Cry1 hybrid delta-endotoxin
CC gene.
XX
SQ Sequence 20 BP; 6 A; 6 C; 2 G; 6 T; 0 other;

Query Match 100.0%; Score 8; DB 19; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.8e+03;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGTTCG 8
||||||||
Db 17 GACGTTCG 10

RESULT 20
AAV21458/c
ID AAV21458 standard; DNA; 20 BP.
XX
AC AAV21458;
XX
DT 20-JUL-1998 (first entry)
XX
DE Mycobacterium tuberculosis sequence probe 3.
XX
KW MTB; amplification; phosphate; inhibitor; Ligase chain reaction; ss.
XX
OS Synthetic.
OS Mycobacterium tuberculosis.
XX
PN WO9806877-A2.
XX
PD 19-FEB-1998.
XX
PF 13-AUG-1997; 97WO-US14380.
XX
PR 15-AUG-1996; 96US-0698573.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Erickson D, Halaka FG, He Q, Leckie GW, Lin B;
XX
DR WPI; 1998-159562/14.
XX
PT Removing inhibitors from nucleic acid amplification reactions - by
PT acidifying test sample containing target sequence and replacing
PT liquid with buffer for amplification
XX
PS Examples; Page 13; 16pp; English.
XX
CC Probes AAV21456-V21459 were used in the amplification of a Mycobacterium
CC tuberculosis (MTB) gene sequence, to demonstrate a method for
CC alleviating inhibition in nucleic acid amplification assays. This method
CC is used for removing amplification inhibitors such as phosphate
CC inhibitors from test samples such as sputum. The MTB gene sequence was
CC successfully amplified and detected using this method to prepare the
CC sample for gap LCR.
XX
SQ Sequence 20 BP; 6 A; 5 C; 6 G; 3 T; 0 other;

Query Match 100.0%; Score 8; DB 19; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.8e+03;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGTTCG 8
||||||||
Db 18 GACGTTCG 11

RESULT 21
AAZ06010
ID AAZ06010 standard; DNA; 20 BP.
XX
AC AAZ06010;
XX
DT 07-OCT-1999 (first entry)
XX
DE PCR primer used to amplify an ORF of Chlamydia trachomatis.
XX
KW Vaccine; eye disease; conventional trachoma; nonendemic trachoma;
KW paratrachoma; inclusion conjunctivitis; genital disease; perihepatitis;
KW nongonococcal uretritis; epidymitis; cervicitis; salpingitis; PCR primer;
KW bartholinitis; pneumopathy; venereal lymphogranulomatosis; ss.
XX
OS Synthetic.
OS Chlamydia trachomatis.
XX
PN WO9928475-A2.
XX
PD 10-JUN-1999.
XX
PF 27-NOV-1998; 98WO-IB01939.
XX
PR 04-NOV-1998; 98US-0107077.
PR 28-NOV-1997; 97FR-0015041.
PR 17-DEC-1997; 97FR-0016034.
XX
PA (GEST) GENSET.
XX
PI Griffais R;
XX
DR WPI; 1999-371125/31.
XX
PT Genome sequence of Chlamydia trachomatis
XX
PS Disclosure; Page 1817; 1755pp; English.
XX
CC PCR primers AAZ01426-Z06209 were used to amplify open reading frames
CC (ORFs) of the genome of Chlamydia trachomatis (see AAZ01425). These ORFs
CC encode polypeptides (see AAY36754-Y37949) which can be used as vaccines
CC against Chlamydia trachomatis. Antisense and ribozyme sequences
CC can also be used to control growth of the microorganism. Chlamydia
CC trachomatis is responsible for a large number of diseases, e.g. eye
CC diseases such as conventional trachoma, nonendemic trachoma,
CC paratrachoma, and inclusion conjunctivitis; genital diseases such as
CC nongonococcal uretritis, epidymitis, cervicitis, salpingitis,
CC perihepatitis, bartholinitis; pneumopathy in breast feeding infants;
CC and venereal lymphogranulomatosis. The polypeptides of the
CC invention may be of use in treating these diseases.
XX
SQ Sequence 20 BP; 4 A; 3 C; 7 G; 6 T; 0 other;

Query Match 100.0%; Score 8; DB 20; Length 20;

```
  Best Local Similarity   100.0%;  Pred. No. 7.8e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db       3 gacgttcg 10


RESULT  22
AAX93946
ID   AAX93946 standard; DNA; 20 BP.
XX
AC   AAX93946;
XX
DT   13-SEP-1999  (first entry)
XX
DE   PCR primer used to amplify an ORF of Chlamydia pneumoniae.
XX
KW   Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KW   sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
KW   vaccine; neutralising epitope; PCR primer; ss.
XX
OS   Synthetic.
OS   Chlamydia pneumoniae.
XX
PN   WO9927105-A2.
XX
PD   03-JUN-1999.
XX
PF   20-NOV-1998;   98WO-IB01890.
XX
PR   04-NOV-1998;   98US-0107078.
PR   21-NOV-1997;   97FR-0014673.
XX
PA   (GEST ) GENSET.
XX
PI   Griffais R;
XX
DR   WPI; 1999-357842/30.
XX
PT   Genome sequence of Chlamydia pneumoniae
XX
PS   Page 1631; Disclosure; 1912pp; English.
XX
CC   AAX91991-X97517 represent PCR primers used to amplify open reading
CC   frames and other nucleic acid sequences from the genome of
CC   Chlamydia pneumoniae (see AAX91990). C. pneumoniae causes respiratory
CC   disease such as pneumonia and bronchitis and is thought to be a
CC   contributing factor in heart disease, sarcoidosis, sinusitis, purulent
CC   otitis media, erythema nodosum or pharyngitis. The polypeptides encoded
CC   by the open reading frames of the C. pneumoniae genome (see AAY34584-
CC   AAY35879) can be used in immunogenic compositions as vaccines. Vectors
CC   containing C. pneumoniae nucleotides sequences can also be used as
CC   immunogenic compositions, especially where the vector directs the
CC   expression of a neutralising epitope of C. pneumoniae.
XX
SQ   Sequence 20 BP; 4 A; 4 C; 6 G; 6 T; 0 other;


  Query Match             100.0%;  Score 8;  DB 20;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db       1 gacgttcg 8


RESULT  23
AAV80112
ID   AAV80112 standard; DNA; 20 BP.
XX
AC   AAV80112;
XX
DT   12-MAR-1999  (first entry)
XX
DE   Oligo used in experiments for stimulation of cytokine production.
XX
KW   Immunomodulatory; immunostimulatory; octanucleotide; immune regulation;
KW   ISS: cancer; allergy; asthma; hepatitis B infection; papillomavirus;
KW   human immunodeficiency virus; influenza; herpes; M. tuberculosis; ss;
KW   B. pertussis; malaria; plasmodia; Leishmania; Trypanosoma; Schistosoma.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   modified_base   8
FT                   /*tag=  a
FT                   /note= "5-bromocytosine"
XX
PN   WO9855495-A2.
XX
PD   10-DEC-1998.
XX
PF   05-JUN-1998;   98WO-US11578.
XX
PR   06-JUN-1997;   97US-0048793.
XX
PA   (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI   Dina D,  Roman M,  Schwartz D;
XX
DR   WPI; 1999-059898/05.
XX
PT   Immunostimulatory oligonucleotides regulate the immune system - and
PT   contain an immune-stimulating octanucleotide sequence; for treating
PT   cancer, allergic and infectious diseases
XX
PS   Example 2; Page 30; 63pp; English.
XX
CC   The invention relates to immunomodulatory oligonucleotides that comprise
CC   at least 1 immunostimulatory octanucleotide sequence (ISS) where the ISS
CC   sequences are selected from the group consisting of AACGTTCC, AACGTTCG,
CC   GACGTTCC, and GACGTTCG. The immunomodulatory sequences are used to treat
CC   patients needing immune regulation, such as those suffering from cancer,
CC   an allergic disease and asthma. They are also used to prevent infectious
CC   diseases such as influenza, herpes, hepatitis B, human immunodeficiency
CC   and papillomavirus, Hemophilus influenza, Mycobacterium tuberculosis and
CC   Bordetella pertussis, malarial plasmodia, Leishmania, Trypanosoma and
CC   Schistosoma. The immunomodulatory sequences are used to screen for human
CC   immunostimulatory activity by incubating macrophage cells and the
CC   oligonucleotide; and determining the relative amount of Th1-biased
CC   cytokines in the supernatant. Sequences AAV80104 to AAV80116 represent
CC   oligonucleotides that were tested for immunostimulatory activity. These
CC   were used in experiments for the stimulation of cytokine production and
CC   were found to lack immunostimulatory activity. The invention provides
CC   specific claimed examples (AAV80096-103) of immunomodulatory sequences.
XX
SQ   Sequence 20 BP; 3 A; 5 C; 5 G; 7 T; 0 other;


  Query Match             100.0%;  Score 8;  DB 20;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db       6 gacgttcg 13


RESULT  24
AAH23258/c
ID   AAH23258 standard; DNA; 20 BP.
XX
AC   AAH23258;
XX
```

```
DT   17-SEP-2001  (first entry)
XX
DE   Human MMIF mRNA inhibiting antisense oligo ISIS #115632.
XX
KW   Macrophage migration inhibitory factor; MMIF; antisense; neurological;
KW   hyperproliferation; nootropic; antihormonal; immunosuppresive; human;
KW   antiinflammatory; cytostatic; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO200153317-A1.
XX
PD   26-JUL-2001.
XX
PF   16-JAN-2001; 2001WO-US01475.
XX
PR   20-JAN-2000; 2000US-0489869.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Murray SF,  Cowsert LM,  Wyatt JR;
XX
DR   WPI; 2001-451899/48.
XX
PT   New antisense compound(s) are useful to inhibit a nucleic acid molecule
PT   encoding macrophage migration inhibitory factor -
XX
PS   Example 15; Page 83; 105pp; English.
XX
CC   The invention relates to antisense oligonucleotides 8-30 nucleotides in
CC   length targeted to a nucleic acid molecule encoding macrophage migration
CC   inhibitory factor (MMIF), where the antisense compound specifically
CC   hybridizes with and inhibits the expression of MMIF. The antisense
CC   nucleotides are useful for the treatment of a disease or condition
CC   associated with MMIF such as neurological, hormonal, immune, inflammatory
CC   or hyperproliferative disorder. Sequences AAH23191-268 represent chimeric
CC   antisense phosphorothioate oligonucleotides used for inhibition of human
CC   MMIF mRNA expression.
XX
SQ   Sequence 20 BP; 3 A; 9 C; 5 G; 3 T; 0 other;


  Query Match             100.0%;  Score 8;  DB 22;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
            ||||||||
Db      13 GACGTTCG 6


RESULT  25
AAF99392
ID   AAF99392 standard; DNA; 20 BP.
XX
AC   AAF99392;
XX
DT   12-JUN-2001  (first entry)
XX
DE   Immunostimulatory nucleic acid #508.
XX
KW   Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
KW   immunostimulatory; tumour; viral infection; bacterial infection;
KW   fungal infection; parasitic infection; cancer; asthma;
KW   infectious disease; allergy; immune deficiency; phosphorothioate; ss.
XX
OS   Synthetic.
XX
PN   WO200122972-A2.
XX
PD   05-APR-2001.
XX
PF   25-SEP-2000; 2000WO-US26383.
XX
PR   25-SEP-1999;   99US-0156113.
PR   27-SEP-1999;   99US-0156135.
PR   23-AUG-2000; 2000US-0227436.
XX
PA   (IOWA ) UNIV IOWA RES FOUND.
PA   (COLE-) COLEY PHARM GMBH.
XX
PI   Krieg AM,  Schetter C,  Vollmer J;
XX
DR   WPI; 2001-273485/28.
XX
PT   Vaccinating against tumors, infectious diseases, allergies and asthma
PT   using immunostimulatory Py-rich and TG nucleic acids -
XX
PS   Claim 101; Page 48; 338pp; English.
XX
CC   The present invention relates to a method for stimulating an immune
CC   response. The method comprises administering an immunostimulatory nucleic
CC   acid to a non-rodent subject in sufficient quantity to stimulate an
CC   immune response. The present sequence is one such immunostimulatory
CC   nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
CC   (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
CC   against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
CC   and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
CC   haemophilus, campylobacter, clostridium, Escherichia coli and/or
CC   staphylococcus), fungal antigens and/or parasitic antigens. The method is
CC   also useful for preventing cancer, asthma, infectious disease, allergy or
CC   immune deficiency. The present sequence can also be used to redirect a
CC   Th2 to a Th1 immune response and to activate immune cells.
CC   Note: the present sequence may have a phosphorothioate backbone.
XX
SQ   Sequence 20 BP; 2 A; 3 C; 12 G; 3 T; 0 other;


  Query Match             100.0%;  Score 8;  DB 22;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
            ||||||||
Db       8 gacgttcg 15


RESULT  26
AAF99395
ID   AAF99395 standard; DNA; 20 BP.
XX
AC   AAF99395;
XX
DT   12-JUN-2001  (first entry)
XX
DE   Immunostimulatory nucleic acid #511.
XX
KW   Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
KW   immunostimulatory; tumour; viral infection; bacterial infection;
KW   fungal infection; parasitic infection; cancer; asthma;
KW   infectious disease; allergy; immune deficiency; phosphorothioate; ss.
XX
OS   Synthetic.
XX
PN   WO200122972-A2.
XX
PD   05-APR-2001.
XX
PF   25-SEP-2000; 2000WO-US26383.
XX
PR   25-SEP-1999;   99US-0156113.
PR   27-SEP-1999;   99US-0156135.
PR   23-AUG-2000; 2000US-0227436.
XX
PA   (IOWA ) UNIV IOWA RES FOUND.
```

PA (COLE-) COLEY PHARM GMBH.
XX
PI Krieg AM, Schetter C, Vollmer J;
XX
DR WPI; 2001-273485/28.
XX
PT Vaccinating against tumors, infectious diseases, allergies and asthma
PT using immunostimulatory Py-rich and TG nucleic acids -
XX
PS Claim 101; Page 48; 338pp; English.
XX
CC The present invention relates to a method for stimulating an immune
CC response. The method comprises administering an immunostimulatory nucleic
CC acid to a non-rodent subject in sufficient quantity to stimulate an
CC immune response. The present sequence is one such immunostimulatory
CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
CC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
CC haemophilus, campylobacter, clostridium, Escherichia coli and/or
CC staphylococcus), fungal antigens and/or parasitic antigens. The method is
CC also useful for preventing cancer, asthma, infectious disease, allergy or
CC immune deficiency. The present sequence can also be used to redirect a
CC Th2 to a Th1 immune response and to activate immune cells.
CC Note: the present sequence may have a phosphorothioate backbone.
XX
SQ Sequence 20 BP; 2 A; 3 C; 12 G; 3 T; 0 other;

```
  Query Match             100.0%;  Score 8;  DB 22;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db       8 gacgttcg 15
```

RESULT 27
AAF99397
ID AAF99397 standard; DNA; 20 BP.
XX
AC AAF99397;
XX
DT 12-JUN-2001 (first entry)
XX
DE Immunostimulatory nucleic acid #513.
XX
KW Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
KW immunostimulatory; tumour; viral infection; bacterial infection;
KW fungal infection; parasitic infection; cancer; asthma;
KW infectious disease; allergy; immune deficiency; phosphorothioate; ss.
XX
OS Synthetic.
XX
PN WO200122972-A2.
XX
PD 05-APR-2001.
XX
PF 25-SEP-2000; 2000WO-US26383.
XX
PR 25-SEP-1999; 99US-0156113.
PR 27-SEP-1999; 99US-0156135.
PR 23-AUG-2000; 2000US-0227436.
XX
PA (IOWA) UNIV IOWA RES FOUND.
PA (COLE-) COLEY PHARM GMBH.
XX
PI Krieg AM, Schetter C, Vollmer J;
XX
DR WPI; 2001-273485/28.
XX
PT Vaccinating against tumors, infectious diseases, allergies and asthma
PT using immunostimulatory Py-rich and TG nucleic acids -
XX
PS Claim 101; Page 48; 338pp; English.
XX
CC The present invention relates to a method for stimulating an immune
CC response. The method comprises administering an immunostimulatory nucleic
CC acid to a non-rodent subject in sufficient quantity to stimulate an
CC immune response. The present sequence is one such immunostimulatory
CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
CC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
CC haemophilus, campylobacter, clostridium, Escherichia coli and/or
CC staphylococcus), fungal antigens and/or parasitic antigens. The method is
CC also useful for preventing cancer, asthma, infectious disease, allergy or
CC immune deficiency. The present sequence can also be used to redirect a
CC Th2 to a Th1 immune response and to activate immune cells.
CC Note: the present sequence may have a phosphorothioate backbone.
XX
SQ Sequence 20 BP; 6 A; 3 C; 8 G; 3 T; 0 other;

```
  Query Match             100.0%;  Score 8;  DB 22;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db       8 gacgttcg 15
```

RESULT 28
AAF99651
ID AAF99651 standard; DNA; 20 BP.
XX
AC AAF99651;
XX
DT 12-JUN-2001 (first entry)
XX
DE Immunostimulatory nucleic acid #767.
XX
KW Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
KW immunostimulatory; tumour; viral infection; bacterial infection;
KW fungal infection; parasitic infection; cancer; asthma;
KW infectious disease; allergy; immune deficiency; phosphorothioate; ss.
XX
OS Synthetic.
XX
PN WO200122972-A2.
XX
PD 05-APR-2001.
XX
PF 25-SEP-2000; 2000WO-US26383.
XX
PR 25-SEP-1999; 99US-0156113.
PR 27-SEP-1999; 99US-0156135.
PR 23-AUG-2000; 2000US-0227436.
XX
PA (IOWA) UNIV IOWA RES FOUND.
PA (COLE-) COLEY PHARM GMBH.
XX
PI Krieg AM, Schetter C, Vollmer J;
XX
DR WPI; 2001-273485/28.
XX
PT Vaccinating against tumors, infectious diseases, allergies and asthma
PT using immunostimulatory Py-rich and TG nucleic acids -
XX
PS Claim 101; Page 55; 338pp; English.
XX
CC The present invention relates to a method for stimulating an immune
CC response. The method comprises administering an immunostimulatory nucleic
CC acid to a non-rodent subject in sufficient quantity to stimulate an

```
CC   immune response. The present sequence is one such immunostimulatory
CC   nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
CC   (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
CC   against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
CC   and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
CC   haemophilus, campylobacter, clostridium, Escherichia coli and/or
CC   staphylococcus), fungal antigens and/or parasitic antigens. The method is
CC   also useful for preventing cancer, asthma, infectious disease, allergy or
CC   immune deficiency. The present sequence can also be used to redirect a
CC   Th2 to a Th1 immune response and to activate immune cells.
CC   Note: the present sequence may have a phosphorothioate backbone.
XX
SQ   Sequence 20 BP; 6 A; 3 C; 8 G; 3 T; 0 other;

  Query Match             100.0%;  Score 8;  DB 22;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db       8 gacgttcg 15


RESULT  29
AAF99692
ID   AAF99692 standard; DNA; 20 BP.
XX
AC   AAF99692;
XX
DT   12-JUN-2001  (first entry)
XX
DE   Immunostimulatory nucleic acid #808.
XX
KW   Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
KW   immunostimulatory; tumour; viral infection; bacterial infection;
KW   fungal infection; parasitic infection; cancer; asthma;
KW   infectious disease; allergy; immune deficiency; phosphorothioate; ss.
XX
OS   Synthetic.
XX
PN   WO200122972-A2.
XX
PD   05-APR-2001.
XX
PF   25-SEP-2000; 2000WO-US26383.
XX
PR   25-SEP-1999;   99US-0156113.
PR   27-SEP-1999;   99US-0156135.
PR   23-AUG-2000;  2000US-0227436.
XX
PA   (IOWA ) UNIV IOWA RES FOUND.
PA   (COLE-) COLEY PHARM GMBH.
XX
PI   Krieg AM,  Schetter C,  Vollmer J;
XX
DR   WPI; 2001-273485/28.
XX
PT   Vaccinating against tumors, infectious diseases, allergies and asthma
PT   using immunostimulatory Py-rich and TG nucleic acids -
XX
PS   Claim 101; Page 56; 338pp; English.
XX
CC   The present invention relates to a method for stimulating an immune
CC   response. The method comprises administering an immunostimulatory nucleic
CC   acid to a non-rodent subject in sufficient quantity to stimulate an
CC   immune response. The present sequence is one such immunostimulatory
CC   nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
CC   (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
CC   against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
CC   and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
CC   haemophilus, campylobacter, clostridium, Escherichia coli and/or
CC   staphylococcus), fungal antigens and/or parasitic antigens. The method is
CC   also useful for preventing cancer, asthma, infectious disease, allergy or
CC   immune deficiency. The present sequence can also be used to redirect a
CC   Th2 to a Th1 immune response and to activate immune cells.
CC   Note: the present sequence may have a phosphorothioate backbone.
XX
SQ   Sequence 20 BP; 1 A; 9 C; 8 G; 2 T; 0 other;

  Query Match             100.0%;  Score 8;  DB 22;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db       3 gacgttcg 10


RESULT  30
AAF60653/c
ID   AAF60653 standard; DNA; 20 BP.
XX
AC   AAF60653;
XX
DT   30-APR-2001  (first entry)
XX
DE   HSV-1 R15 PCR primer #1.
XX
KW   Antiviral; antitumour; cytostatic; vaccine; PCR primer;
KW   unique small 8 gene; gE protein; unique small 12 gene; ICP47 protein;
KW   astrocytomam brain tumour; ss.
XX
OS   Herpes simplex virus.
XX
PN   WO200109361-A1.
XX
PD   08-FEB-2001.
XX
PF   03-AUG-2000; 2000WO-IL00469.
XX
PR   03-AUG-1999;   99IL-0131212.
XX
PA   (YISS ) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
PA   (HADA-) HADASIT MEDICAL RES SERVICES & DEV.
XX
PI   Becker Y,  Asher Y,  Bujanover S,  Ben-Hur T,  Darai G;
XX
DR   WPI; 2001-182967/18.
XX
PT   Recombinant herpes simplex virus-1, HSV-1 R15, comprising a mutated
PT   genome with mutations or deletions in unique small 8 and 12 genes,
PT   useful as components of anti-herpes vaccines -
XX
PS   Disclosure; Page 24; 81pp; English.
XX
CC   The present invention relates to a recombinant herpes simplex virus (HSV)
CC   (HSV-1 R15; AAF60648), comprising a mutant genome of HSV-1, with a
CC   deletion or mutation in a unique small 8 gene region resulting either in
CC   expression of a non-functional or no expression of a gE protein, and a
CC   deletion or mutation in a unique small 12 gene region resulting either in
CC   expression of a non-functional or no expression of an ICP47 protein.
CC   HSV-1 R15 is useful as an effective carrier of genes coding for an
CC   immunogenic protein of a desired virus for the purpose of vaccination
CC   against the virus. The property of HSV-1 R15 recombinant to be able to
CC   replicate in the brain astrocyte cell culture, while being incapable of
CC   affecting nerve cells, can be used to selectively transfect only tumour
CC   astrocytes (astrocytoma) while leaving nerve cells uninfected. This
CC   property may be used to cure astrocytoma brain tumours, either by simply
CC   injecting HSV-1 R15 of the invention stereotactically to the tumour so
CC   its natural replication will destroy the tumour, or inserting the genome
CC   of the recombinant virus. HSV-1 R15 can also be used to prepare a vaccine
CC   and is useful for immunising a subject against HSV-1 or any of herpes
CC   viruses 2-8. The present sequence is a PCR primer for HSV-1 R15.
XX
```

```
SQ   Sequence 20 BP; 5 A; 8 C; 3 G; 4 T; 0 other;

  Query Match             100.0%;  Score 8;  DB 22;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      14 GACGTTCG 7


RESULT  31
AAF60804/c
ID   AAF60804 standard; DNA; 20 BP.
XX
AC   AAF60804;
XX
DT   04-MAY-2001  (first entry)
XX
DE   S. cerevisiae MET16 PCR primer SEQ ID 6.
XX
KW   Microorganism; sulfite production; sulfite cycle; food production; wine;
KW   beer; taste stabilizer; non-volatile complex; carbonyl compound; yeast;
KW   oxidation; fermentation; PCR primer; ss.
XX
OS   Saccharomyces cerevisiae.
XX
PN   DE19923950-A1.
XX
PD   25-JAN-2001.
XX
PF   25-MAY-1999;   99DE-1023950.
XX
PR   25-MAY-1999;   99DE-1023950.
XX
PA   (STAH/) STAHL U.
XX
PI   Donalis U,  Stahl U;
XX
DR   WPI; 2001-148153/16.
XX
PT   New microorganisms that produce high sulfite levels at a late stage in
PT   their growth, useful for producing beer, prevent development of
PT   off-flavors by oxidation -
XX
PS   Example 2; Page 7; 14pp; German.
XX
CC   This invention describes novel microorganisms (A) able to produce delayed
CC   and large amounts of sulfite. The microorganisms comprise a DNA construct
CC   (I) containing one or more genes (II) involved in the sulfite cycle under
CC   the control of a promoter. The high sulfite concentration appears at a
CC   late stage of substrate utilization, in the stationary growth phase, in
CC   the last third of the exponential growth phase or at a cell density of
CC   60-90% of that achieved in the growth phase. (A), particularly bacteria
CC   and/or yeast, are used for the production of foods, wine, beer or desired
CC   metabolic end products. Sulfite stabilizes the taste of beer by forming
CC   non-volatile complexes with carbonyl compounds (formed by oxidation and
CC   responsible for off-flavors) and by preventing oxidation (reducing
CC   agent). (A) produce significant amounts of sulfite only at a late stage
CC   in its growth, after the fermentation product has formed, avoiding
CC   premature formation of complexes and eliminating the need to add sulfite
CC   to the finished product.
XX
SQ   Sequence 20 BP; 4 A; 7 C; 3 G; 6 T; 0 other;

  Query Match             100.0%;  Score 8;  DB 22;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      19 GACGTTCG 12


RESULT  32
AAX34275/c
ID   AAX34275 standard; DNA; 21 BP.
XX
AC   AAX34275;
XX
DT   06-JUL-1999  (first entry)
XX
DE   Primer AsdF2 for P.aeruginosa Asd gene.
XX
KW   Identification; genome; insertional mutagenesis; amplification; primer;
KW   PCR; essential gene test; mutation; phenotype; FtsZ; Potato blight virus;
KW   equine encephalitis virus; HIV; influenza virus; herpes virus; fungus;
KW   cytomegalovirus; yeast; Candida; Cryptococcus; Histoplasma; Blastomyces;
KW   Coccioides; Aspergillus; Fusarium; Trychophyton; ss.
XX
OS   Synthetic.
OS   Pseudomonas aeruginosa.
XX
PN   WO9915644-A2.
XX
PD   01-APR-1999.
XX
PF   21-SEP-1998;   98WO-CA00893.
XX
PR   19-SEP-1997;   97CA-2215870.
XX
PA   (UYLA-) UNIV LAVAL.
XX
PI   Cardinal G,  Levesque RC,  Sanschagrin F;
XX
DR   WPI; 1999-254705/21.
XX
PT   Identification of essential genes in a genome of e.g. Herpes virus
XX
PS   Example 2; Page 30; 45pp; English.
XX
CC   The invention relates to a method of identifying essential and
CC   non-essential genes in a chosen genome, based on insertional mutagenesis
CC   of a population of cells or DNA molecules, and subsequent amplification.
CC   The method is designated the essential gene test (EGT), and is based on
CC   the premise that a mutation inactivating an essential gene should give
CC   rise, in vivo, to a lethal phenotype.  Primers AAX34275-X34276 were used
CC   to PCR amplify a fragment of the Asd gene from Pseudomonas aeruginosa
CC   strain PAO1.  The methods can be used to identify essential genes in
CC   disease causing organisms such as viruses (e.g. Potato blight virus,
CC   equine encephalitis virus, Human immunodeficiency virus, Influenza
CC   virus, Herpes virus, and Cytomegalovirus), and in single eukaryotic
CC   cells of fungi and yeast causing diseases such as mycoses (e.g. Candida,
CC   Cryptococcus, Histoplasma, Blastomyces, Coccioides, Aspergillus,
CC   Fusarium and Trychophyton).
XX
SQ   Sequence 21 BP; 4 A; 7 C; 6 G; 4 T; 0 other;

  Query Match             100.0%;  Score 8;  DB 20;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 7.7e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      17 GACGTTCG 10


RESULT  33
AAF98750
ID   AAF98750 standard; DNA; 22 BP.
XX
AC   AAF98750;
XX
```

```
DT   11-JUN-2001  (first entry)
XX
DE   Human IFN-alpha immunostimulatory nucleic acid SEQ ID NO: 20.
XX
KW   Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
KW   viral infection; phosphorothioate backbone; palindrome; cancer; ds.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   modified_base   1..2
FT                   /*tag= a
FT                   /mod_base= "OTHER"
FT                   /note= "phosphorothioate linkage"
FT   modified_base   17..21
FT                   /*tag= b
FT                   /mod_base= "OTHER"
FT                   /note= "phosphorothioate linkage"
XX
PN   WO200122990-A2.
XX
PD   05-APR-2001.
XX
PF   27-SEP-2000; 2000WO-US26527.
XX
PR   27-SEP-1999;   99US-0156147.
XX
PA   (COLE-) COLEY PHARM GROUP INC.
PA   (IOWA ) UNIV IOWA RES FOUND.
XX
PI   Hartmann G,  Bratzler RL,  Krieg A;
XX
DR   WPI; 2001-290487/30.
XX
PT   Improving the efficacy of treatments involving the administration of
PT   interferon-alpha by co-administering an isolated immunostimulatory
PT   nucleic acid -
XX
PS   Claim 201; Page 103; 168pp; English.
XX
CC   The present invention describes an improvement to a method requiring the
CC   administration of interferon alpha (IFN-alpha), involving administering
CC   an immunostimulatory nucleic acid (ISNA). The sequences of a number of
CC   such nucleic acids are also provided. These may comprise oligonucleotides
CC   with phosphorothioate backbones, palindromes, or G-rich sequences. The
CC   sequences of the invention are useful in the treatment of proliferative
CC   diseases, such as cancers, and viral infections. The present sequence is
CC   an example of an immunostimulatory oligonucleotide.
XX
SQ   Sequence 22 BP; 3 A; 3 C; 13 G; 3 T; 0 other;

  Query Match             100.0%;  Score 8;  DB 22;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 7.7e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db       4 gacgttcg 11


RESULT  34
AAF98873
ID   AAF98873 standard; DNA; 22 BP.
XX
AC   AAF98873;
XX
DT   11-JUN-2001  (first entry)
XX
DE   Immunostimulatory nucleic acid assay control oligo SEQ ID NO: 154.
XX
KW   Immunostimulatory nucleic acid; ISNA; human; interferon alpha; IFN-alpha;
KW   viral infection; phosphorothioate backbone; palindrome; cancer; ds.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   modified_base   1..22
FT                   /*tag= a
FT                   /mod_base= "OTHER"
FT                   /note= "phosphorothioate linkage"
XX
PN   WO200122990-A2.
XX
PD   05-APR-2001.
XX
PF   27-SEP-2000; 2000WO-US26527.
XX
PR   27-SEP-1999;   99US-0156147.
XX
PA   (COLE-) COLEY PHARM GROUP INC.
PA   (IOWA ) UNIV IOWA RES FOUND.
XX
PI   Hartmann G,  Bratzler RL,  Krieg A;
XX
DR   WPI; 2001-290487/30.
XX
PT   Improving the efficacy of treatments involving the administration of
PT   interferon-alpha by co-administering an isolated immunostimulatory
PT   nucleic acid -
XX
PS   Example 17; Page 163; 168pp; English.
XX
CC   The present invention describes an improvement to a method requiring the
CC   administration of interferon alpha (IFN-alpha), involving administering
CC   an immunostimulatory nucleic acid (ISNA). The sequences of a number of
CC   such nucleic acids are also provided. These may comprise oligonucleotides
CC   with phosphorothioate backbones, palindromes, or G-rich sequences. The
CC   sequences of the invention are useful in the treatment of proliferative
CC   diseases, such as cancers, and viral infections. The present sequence is
CC   an example of an immunostimulatory oligonucleotide.
XX
SQ   Sequence 22 BP; 3 A; 3 C; 13 G; 3 T; 0 other;

  Query Match             100.0%;  Score 8;  DB 22;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 7.7e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db       4 gacgttcg 11


RESULT  35
AAF99680
ID   AAF99680 standard; DNA; 22 BP.
XX
AC   AAF99680;
XX
DT   12-JUN-2001  (first entry)
XX
DE   Immunostimulatory nucleic acid #796.
XX
KW   Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
KW   immunostimulatory; tumour; viral infection; bacterial infection;
KW   fungal infection; parasitic infection; cancer; asthma;
KW   infectious disease; allergy; immune deficiency; phosphorothioate; ss.
XX
OS   Synthetic.
XX
PN   WO200122972-A2.
XX
PD   05-APR-2001.
XX
PF   25-SEP-2000; 2000WO-US26383.
```

```
XX
PR   25-SEP-1999;   99US-0156113.
PR   27-SEP-1999;   99US-0156135.
PR   23-AUG-2000; 2000US-0227436.
XX
PA   (IOWA ) UNIV IOWA RES FOUND.
PA   (COLE-) COLEY PHARM GMBH.
XX
PI   Krieg AM,  Schetter C,  Vollmer J;
XX
DR   WPI; 2001-273485/28.
XX
PT   Vaccinating against tumors, infectious diseases, allergies and asthma
PT   using immunostimulatory Py-rich and TG nucleic acids -
XX
PS   Claim 101; Page 55; 338pp; English.
XX
CC   The present invention relates to a method for stimulating an immune
CC   response. The method comprises administering an immunostimulatory nucleic
CC   acid to a non-rodent subject in sufficient quantity to stimulate an
CC   immune response. The present sequence is one such immunostimulatory
CC   nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
CC   (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
CC   against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
CC   and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
CC   haemophilus, campylobacter, clostridium, Escherichia coli and/or
CC   staphylococcus), fungal antigens and/or parasitic antigens. The method is
CC   also useful for preventing cancer, asthma, infectious disease, allergy or
CC   immune deficiency. The present sequence can also be used to redirect a
CC   Th2 to a Th1 immune response and to activate immune cells.
CC   Note: the present sequence may have a phosphorothioate backbone.
XX
SQ   Sequence 22 BP; 1 A; 10 C; 7 G; 4 T; 0 other;

  Query Match             100.0%;  Score 8;  DB 22;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 7.7e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db       5 gacgttcg 12


RESULT  36
AAF99776
ID   AAF99776 standard; DNA; 22 BP.
XX
AC   AAF99776;
XX
DT   12-JUN-2001  (first entry)
XX
DE   Immunostimulatory nucleic acid #892.
XX
KW   Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
KW   immunostimulatory; tumour; viral infection; bacterial infection;
KW   fungal infection; parasitic infection; cancer; asthma;
KW   infectious disease; allergy; immune deficiency; phosphorothioate; ss.
XX
OS   Synthetic.
XX
PN   WO200122972-A2.
XX
PD   05-APR-2001.
XX
PF   25-SEP-2000; 2000WO-US26383.
XX
PR   25-SEP-1999;   99US-0156113.
PR   27-SEP-1999;   99US-0156135.
PR   23-AUG-2000; 2000US-0227436.
XX
PA   (IOWA ) UNIV IOWA RES FOUND.
PA   (COLE-) COLEY PHARM GMBH.
XX
PI   Krieg AM,  Schetter C,  Vollmer J;
XX
DR   WPI; 2001-273485/28.
XX
PT   Vaccinating against tumors, infectious diseases, allergies and asthma
PT   using immunostimulatory Py-rich and TG nucleic acids -
XX
PS   Claim 101; Page 57; 338pp; English.
XX
CC   The present invention relates to a method for stimulating an immune
CC   response. The method comprises administering an immunostimulatory nucleic
CC   acid to a non-rodent subject in sufficient quantity to stimulate an
CC   immune response. The present sequence is one such immunostimulatory
CC   nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
CC   (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
CC   against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
CC   and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
CC   haemophilus, campylobacter, clostridium, Escherichia coli and/or
CC   staphylococcus), fungal antigens and/or parasitic antigens. The method is
CC   also useful for preventing cancer, asthma, infectious disease, allergy or
CC   immune deficiency. The present sequence can also be used to redirect a
CC   Th2 to a Th1 immune response and to activate immune cells.
CC   Note: the present sequence may have a phosphorothioate backbone.
XX
SQ   Sequence 22 BP; 3 A; 3 C; 13 G; 3 T; 0 other;

  Query Match             100.0%;  Score 8;  DB 22;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 7.7e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db       4 gacgttcg 11


RESULT  37
AAF99832
ID   AAF99832 standard; DNA; 22 BP.
XX
AC   AAF99832;
XX
DT   12-JUN-2001  (first entry)
XX
DE   Immunostimulatory nucleic acid #948.
XX
KW   Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
KW   immunostimulatory; tumour; viral infection; bacterial infection;
KW   fungal infection; parasitic infection; cancer; asthma;
KW   infectious disease; allergy; immune deficiency; phosphorothioate; ss.
XX
OS   Synthetic.
XX
PN   WO200122972-A2.
XX
PD   05-APR-2001.
XX
PF   25-SEP-2000; 2000WO-US26383.
XX
PR   25-SEP-1999;   99US-0156113.
PR   27-SEP-1999;   99US-0156135.
PR   23-AUG-2000; 2000US-0227436.
XX
PA   (IOWA ) UNIV IOWA RES FOUND.
PA   (COLE-) COLEY PHARM GMBH.
XX
PI   Krieg AM,  Schetter C,  Vollmer J;
XX
DR   WPI; 2001-273485/28.
XX
PT   Vaccinating against tumors, infectious diseases, allergies and asthma
PT   using immunostimulatory Py-rich and TG nucleic acids -
```

```
XX
PS   Claim 101; Page 58; 338pp; English.
XX
CC   The present invention relates to a method for stimulating an immune
CC   response. The method comprises administering an immunostimulatory nucleic
CC   acid to a non-rodent subject in sufficient quantity to stimulate an
CC   immune response. The present sequence is one such immunostimulatory
CC   nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
CC   (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
CC   against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
CC   and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
CC   haemophilus, campylobacter, clostridium, Escherichia coli and/or
CC   staphylococcus), fungal antigens and/or parasitic antigens. The method is
CC   also useful for preventing cancer, asthma, infectious disease, allergy or
CC   immune deficiency. The present sequence can also be used to redirect a
CC   Th2 to a Th1 immune response and to activate immune cells.
CC   Note: the present sequence may have a phosphorothioate backbone.
XX
SQ   Sequence 22 BP; 3 A; 3 C; 13 G; 3 T; 0 other;


  Query Match             100.0%;  Score 8;  DB 22;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 7.7e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db       4 gacgttcg 11


RESULT  38
AAT31845
ID   AAT31845 standard; DNA; 23 BP.
XX
AC   AAT31845;
XX
DT   10-JAN-1997  (first entry)
XX
DE   Probe/Primer used for amplifying M. tuberculosis target sequence.
XX
KW   Sonication; probe; primer; Mycobacterium tuberculosis; analysis;
KW   detection; ss.
XX
OS   Synthetic.
XX
PN   WO9619301-A1.
XX
PD   27-JUN-1996.
XX
PF   22-DEC-1995;   95WO-US16810.
XX
PR   21-DEC-1995;   95US-0564995.
PR   22-DEC-1994;   94US-0362640.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Halaka FG;
XX
DR   WPI; 1996-309363/31.
XX
PT   Device for sonicating samples, e.g. body fluid sample - having
PT   electrical wave generator, vibrating element and vibratable member
PT   to generate standing sonic wave
XX
PS   Example 3; Page 26; 35pp; English.
XX
CC   A new device for sonicating a sample comprises an electrical wave
CC   generator, a vibrating element electrically connected to the
CC   electrical wave generator and a vibratable member transcersely
CC   secured to the vibrating element.  The device can be used to lyse
CC   cells such as bacteria and viruses so that the intracellular
CC   components can be analysed.  Four probes/primers (AAT31842-45) were
CC   used to amplify and detect a target sequence (AAT42980) generated from
CC   lysed Mycobacterium tuberculosis cells.  These sequences were
CC   published to illustrate the invention.
XX
SQ   Sequence 23 BP; 3 A; 6 C; 7 G; 7 T; 0 other;


  Query Match             100.0%;  Score 8;  DB 17;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 7.7e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db       3 gacgttcg 10


RESULT  39
AAT28760
ID   AAT28760 standard; DNA; 23 BP.
XX
AC   AAT28760;
XX
DT   11-DEC-1996  (first entry)
XX
DE   Probe #4 for Mycobacterium tuberculosis protein antigen b gene.
XX
KW   Probe; Chlamydia trachomatis; amplification inhibition; cryptic plasmid;
KW   ligase chain reaction; LCR; primer; amplify; Mycobacterium tuberculosis;
KW   protein antigen b; pab; Opa A; Neisseria gonorrhoeae; spermidine;
KW   magnesium chloride; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    23
FT                   /*tag= a
FT                   /note= "conjugated with adamantaneacetic acid
FT                          ("adamantane")"
XX
PN   WO9612824-A2.
XX
PD   02-MAY-1996.
XX
PF   18-OCT-1995;   95WO-US12874.
XX
PR   12-OCT-1995;   95US-0532212.
PR   21-OCT-1994;   94US-0331391.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Davis AH,  Lee EH;
XX
DR   WPI; 1996-230622/23.
XX
PT   Use of spermidine to relieve inhibition of ligase chain reaction -
PT   improves amplification yield and allows reaction to proceed in
PT   presence of non-optimal magnesium chloride concentrations
XX
PS   Example 1; Page 15; 33pp; English.
XX
CC   AAT28753-T28768 represent probes used in the method of the invention, to
CC   relieve amplification inhibition in a ligase chain reaction (LCR).
CC   AAT28757-T28760 represent probes for the protein antigen b (pab) gene of
CC   Mycobacterium tuberculosis.  In the method of the invention, the LCR
CC   amplification inhibition is relieved by forming a reaction mixture
CC   containing a clinical test sample, an amount of spermidine effective to
CC   relieve amplification inhibition, and a composition containing two pairs
CC   of probes.  Each pair of probes used contains a primary probe that is
CC   hybridisable to the target sequence.  The two primary probes used are
CC   hybridisable to adjacent or near adjacent positions on the target
CC   sequence.  The primary probes are then hybridised to the target, and
CC   ligated to form a fused product.  The fused product is then dissociated
CC   from the target.  One or more of the primary or secondary probes used
CC   can be modified at one end, so that it is non-ligatable to the other
```

```
CC   primary or secondary probe.  The presence of spermidine alleviates the
CC   inhibitory effects of certain clinical samples on LCR reactions, and
CC   therefore improves amplification yields.  The presence of spermidine also
CC   allows LCR to proceed in the presence of non-optimal concentrations of
CC   magnesium chloride.  This means that LCR can be carried out in subsequent
CC   assay mixtures that require different concentrations of magnesium
CC   chloride.
XX
SQ   Sequence 23 BP; 3 A; 6 C; 7 G; 7 T; 0 other;


  Query Match             100.0%;  Score 8;  DB 17;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 7.7e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db       3 gacgttcg 10


RESULT  40
AAD07148/c
ID   AAD07148 standard; DNA; 23 BP.
XX
AC   AAD07148;
XX
DT   06-AUG-2001  (first entry)
XX
DE   Back module BM-379 for branched modular primer.
XX
KW   Priming site 19379; PCR; polymerase chain reaction; amplification;
KW   branched modular primer; front module; FM; back module; BM-379; ss.
XX
OS   Bacteriophage lambda.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   modified_base   1
FT                   /*tag=  a
FT                   /mod_base= i
FT   mutation        8
FT                   /*tag=  b
FT                   /mod_base= i
XX
PN   US6235889-B1.
XX
PD   22-MAY-2001.
XX
PF   08-MAR-1999;   99US-0264466.
XX
PR   20-DEC-1991;   91US-0810898.
PR   06-MAY-1997;   97US-0852001.
PR   06-FEB-1995;   95US-0384699.
XX
PA   (UYCH-) UNIV CHICAGO.
XX
PI   Ulanovsky L;
XX
DR   WPI; 2001-366426/38.
XX
PT   New composition comprising front and back oligonucleotide modules, each
PT   module has a stem and an arm segment with varying or constant
PT   sequences, useful for amplifying nucleic acid segments such as in
PT   polymerase chain reaction  -
XX
PS   Disclosure; Column 18; 32pp; English.
XX
CC   The present invention relates to compositions for branched modular
CC   primers used in methods for amplifying a nucleic acid segment. The
CC   branched modular primer comprises of front and back oligonucleotide
CC   modules. The front module (FM) and back module (BM) comprise of a stem
CC   segment having a sequence that is the same from module to module and an
CC   arm segment having a sequence that varies from module to module. The arm
CC   of the back and front modules are annealed to a template which contains
CC   the priming site. These modules are designed for priming sites in lambda
CC   phage DNA. The composition is useful for amplifying a nucleic acid
CC   segment, e.g. by polymerase chain reaction (PCR). The present sequence
CC   is back module BM-379 which is annealed to Bacteriophage lambda reverse
CC   priming site 19379 (template) for constructing a branch modular primer.
XX
SQ   Sequence 23 BP; 5 A; 8 C; 5 G; 3 T; 2 other;


  Query Match             100.0%;  Score 8;  DB 22;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 7.7e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      16 GACGTTCG 9


RESULT  41
AAD07155/c
ID   AAD07155 standard; DNA; 23 BP.
XX
AC   AAD07155;
XX
DT   06-AUG-2001  (first entry)
XX
DE   Back module BM-654 for branched modular primer.
XX
KW   Priming site 19654; PCR; polymerase chain reaction; amplification;
KW   branched modular primer; front module; FM; back module; BM-654; ss.
XX
OS   Bacteriophage lambda.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   modified_base   1
FT                   /*tag=  a
FT                   /mod_base= i
FT   modified_base   8
FT                   /*tag=  b
FT                   /mod_base= i
XX
PN   US6235889-B1.
XX
PD   22-MAY-2001.
XX
PF   08-MAR-1999;   99US-0264466.
XX
PR   20-DEC-1991;   91US-0810898.
PR   06-MAY-1997;   97US-0852001.
PR   06-FEB-1995;   95US-0384699.
XX
PA   (UYCH-) UNIV CHICAGO.
XX
PI   Ulanovsky L;
XX
DR   WPI; 2001-366426/38.
XX
PT   New composition comprising front and back oligonucleotide modules, each
PT   module has a stem and an arm segment with varying or constant
PT   sequences, useful for amplifying nucleic acid segments such as in
PT   polymerase chain reaction  -
XX
PS   Disclosure; Column 18; 32pp; English.
XX
CC   The present invention relates to compositions for branched modular
CC   primers used in methods for amplifying a nucleic acid segment. The
CC   branched modular primer comprises of front and back oligonucleotide
CC   modules. The front module (FM) and back module (BM) comprise of a stem
CC   segment having a sequence that is the same from module to module and an
CC   arm segment having a sequence that varies from module to module. The arm
CC   of the back and front modules are annealed to a template which contains
```

```
CC   the priming site. These modules are designed for priming sites in lambda
CC   phage DNA. The composition is useful for amplifying a nucleic acid
CC   segment, e.g. by polymerase chain reaction (PCR). The present sequence
CC   is back module BM-654 which is annealed to Bacteriophage lambda reverse
CC   priming site 19654 (template) for constructing a branch modular primer.
XX
SQ   Sequence 23 BP; 3 A; 6 C; 8 G; 4 T; 2 other;

  Query Match             100.0%;  Score 8;  DB 22;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 7.7e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      16 GACGTTCG 9


RESULT  42
AAF74942/c
ID   AAF74942 standard; DNA; 23 BP.
XX
AC   AAF74942;
XX
DT   23-MAY-2001  (first entry)
XX
DE   Bacteriophage lambda fragment PCR ampification primer SEQ ID NO:3.
XX
KW   Bacteriophage lambda; PCR primer; amplification; genome mapping;
KW   biomedical research; clinical diagnostic; ss.
XX
OS   Bacteriophage lambda.
OS   Synthetic.
XX
PN   US6197556-B1.
XX
PD   06-MAR-2001.
XX
PF   06-MAY-1997;   97US-0852001.
XX
PR   20-DEC-1991;   91US-0810898.
PR   06-FEB-1995;   95US-0384699.
XX
PA   (UYCH-) UNIV CHICAGO.
XX
PI   Ulanovsky L,  Raja MC;
XX
DR   WPI; 2001-256370/26.
XX
PT   Amplifying a template nucleic acid segment, involves annealing a
PT   combination of several branched and/or covered oligonucleotide modules
PT   selected from a pre-synthesized library, to the template DNA  -
XX
PS   Disclosure; Column 18; 33pp; English.
XX
CC   The present invention describes a method for amplifying a template
CC   nucleic acid segment (I), comprising annealing (I) to a branched primer
CC   having front (FOM) and back oligonucleotide modules with arm segments
CC   complementary to a site in (I), extending the arm of FOM to form an
CC   initial extension strand, annealing the strand to a reverse primer (RP),
CC   extending RP to form second initial extension strand, and amplifying the
CC   second strand. The method can be used for amplifying nucleic acid
CC   segments, useful in genome mapping, biomedical research and clinical
CC   diagnostics. The method eliminates the need for custom primer synthesis
CC   in methods to amplify nucleic acid segments. The modular combination of
CC   just a few oligonucleotides essentially mimics the performance of a
CC   conventional, custom-made primer by matching a sequence of a priming
CC   site in the template. AAF74940 to AAF74979 represent oligonucleotide
CC   sequences used in the exemplification of the present invention.
CC   N.B. Any Ns given in the oligonucleotide sequences represent inosine
CC   bases.
XX
SQ   Sequence 23 BP; 5 A; 8 C; 5 G; 3 T; 2 other;

  Query Match             100.0%;  Score 8;  DB 22;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 7.7e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      16 GACGTTCG 9


RESULT  43
AAF74949/c
ID   AAF74949 standard; DNA; 23 BP.
XX
AC   AAF74949;
XX
DT   23-MAY-2001  (first entry)
XX
DE   Bacteriophage lambda fragment PCR ampification primer SEQ ID NO:10.
XX
KW   Bacteriophage lambda; PCR primer; amplification; genome mapping;
KW   biomedical research; clinical diagnostic; ss.
XX
OS   Bacteriophage lambda.
OS   Synthetic.
XX
PN   US6197556-B1.
XX
PD   06-MAR-2001.
XX
PF   06-MAY-1997;   97US-0852001.
XX
PR   20-DEC-1991;   91US-0810898.
PR   06-FEB-1995;   95US-0384699.
XX
PA   (UYCH-) UNIV CHICAGO.
XX
PI   Ulanovsky L,  Raja MC;
XX
DR   WPI; 2001-256370/26.
XX
PT   Amplifying a template nucleic acid segment, involves annealing a
PT   combination of several branched and/or covered oligonucleotide modules
PT   selected from a pre-synthesized library, to the template DNA  -
XX
PS   Disclosure; Column 19; 33pp; English.
XX
CC   The present invention describes a method for amplifying a template
CC   nucleic acid segment (I), comprising annealing (I) to a branched primer
CC   having front (FOM) and back oligonucleotide modules with arm segments
CC   complementary to a site in (I), extending the arm of FOM to form an
CC   initial extension strand, annealing the strand to a reverse primer (RP),
CC   extending RP to form second initial extension strand, and amplifying the
CC   second strand. The method can be used for amplifying nucleic acid
CC   segments, useful in genome mapping, biomedical research and clinical
CC   diagnostics. The method eliminates the need for custom primer synthesis
CC   in methods to amplify nucleic acid segments. The modular combination of
CC   just a few oligonucleotides essentially mimics the performance of a
CC   conventional, custom-made primer by matching a sequence of a priming
CC   site in the template. AAF74940 to AAF74979 represent oligonucleotide
CC   sequences used in the exemplification of the present invention.
CC   N.B. Any Ns given in the oligonucleotide sequences represent inosine
CC   bases.
XX
SQ   Sequence 23 BP; 3 A; 6 C; 8 G; 4 T; 2 other;

  Query Match             100.0%;  Score 8;  DB 22;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 7.7e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
```

```
          ||||||||
Db      16 GACGTTCG 9


RESULT  44
AAQ55886
ID   AAQ55886 standard; DNA; 24 BP.
XX
AC   AAQ55886;
XX
DT   25-JUL-1994  (first entry)
XX
DE   Probe for OriS site I (replication origin segment).
XX
KW   Antiviral; inhibition; replication; therapy; treatment; HSV;
KW   herpes simplex virus; ss.
XX
OS   Synthetic.
XX
PN   CA2068695-A.
XX
PD   15-NOV-1993.
XX
PF   14-MAY-1992;   92CA-2068695.
XX
PR   14-MAY-1992;   92CA-2068695.
XX
PA   (DAND ) DANA FARBER CANCER INST INC.
XX
PI   Amaral CE,  Schaffer PA;
XX
DR   WPI; 1994-043311/06.
XX
PT   Methods for identifying cpds. with antiviral activity - useful
PT   for inhibiting DNA virus replication and treating
PT   virally-infected animals
XX
PS   Disclosure; Figure 7; 54pp; English.
XX
CC   Oligonucleotides (AAQ55874-81) are used in an antiviral composition
CC   which prevents binding of a cellular protein to a specific site on
CC   an origin of replication on the genome of a DNA virus, or
CC   interaction of the protein with an origin binding protein of a DNA
CC   virus.  A number of probes (AAQ55885-55897) were synthesised
CC   specific for a segment of a replication origin of HSV-1 designated
CC   oriS site I.
XX
SQ   Sequence 24 BP; 4 A; 9 C; 5 G; 6 T; 0 other;


  Query Match             100.0%;  Score 8;  DB 15;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 7.7e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db       3 gacgttcg 10


RESULT  45
AAT06577
ID   AAT06577 standard; DNA; 24 BP.
XX
AC   AAT06577;
XX
DT   25-JUN-1996  (first entry)
XX
DE   Probe B' (Set 1) for M. tuberculosis Pab gene nucleotides 347-390.
XX
KW   probe; modified ligase chain reaction; Mycobacterium tuberculosis;
KW   M. avium; M. intracellulare; M. kansasii; detection; diagnosis; ss.
XX
OS   Synthetic.
XX
PN   WO9531571-A2.
XX
PD   23-NOV-1995.
XX
PF   04-MAY-1995;   95WO-US05816.
XX
PR   13-MAY-1994;   94US-0223330.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Kratochvil JD,  Leckie GW,  Odonnell DL,  Solomon NA;
XX
DR   WPI; 1996-010956/01.
XX
PT   New probes for detection of Mycobacterium species - derived from the
PT   16S ribosomal RNA gene, the protein antigen b gene and the 65 kD and
PT   10 kD heat shock protein genes of M.tuberculosis
XX
PS   Example 1; Page 34; 60pp; English.
XX
CC   Probe set 1 (AAT06574-577) was selected to detect a target sequence in
CC   Mycobacterium tuberculosis corresponding to nucleotides 347-390
CC   (AAT06573) of the protein antigen b (pab) gene. The probes were labelled
CC   with carbazole and adamantane. Set 1 was capable of detecting as few as
CC   10 mols. of DNA derived from M. tuberculosis and showed no
CC   cross-reactivity with DNA genomes derived from M. avium, M.
CC   intracellulare and M.kansasii. A modified ligase chain reaction was
CC   utilised which uses two pairs of probes designated A, B (primary probes)
CC   and A', B' (secondary probes). Probe pairs were directed to the same
CC   target strand and ultimately ligated to one another after annealing to
CC   the target strand. At least one of the probes of a pair had a modified
CC   end with respect to the point of ligation.  The modified end had bases
CC   omitted to create a gap between one probe terminus and the next probe
CC   terminus when the pair was annealed to the target sequence.  Other
CC   modified ends include a base mismatched with the target sequence.  The
CC   presence of modified ends reduced the falsely positive signal created by
CC   blunt-end ligation of the complementary probe duplexes to one another in
CC   the absence of target. "Correction" of the modification, in a target
CC   dependent manner, was subsequently carried out to render the probes
CC   ligatable. Once ligated, the fused (reorganised) probe was dissociated
CC   (e.g. melted) from the target and, as with conventional LCR, the process
CC   was repeated for several cycles.
XX
SQ   Sequence 24 BP; 4 A; 6 C; 7 G; 7 T; 0 other;


  Query Match             100.0%;  Score 8;  DB 17;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 7.7e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db       3 gacgttcg 10


Search completed: November 29, 2001, 14:51:05
Job time: 3658 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: November 29, 2001, 14:48:17 ; Search time 64.43 Seconds
(without alignments)
28.121 Million cell updates/sec

Title: FRAG2
Perfect score: 8
Sequence: 1 GACGTTCG 8

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 560984

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	8	100.0	17	4	US-08-485-355B-18	Sequence 18, Appl
c	2	8	100.0	17	4	US-09-025-769B-8	Sequence 8, Appli
c	3	8	100.0	19	4	US-08-348-548-100	Sequence 100, App
c	4	8	100.0	19	5	PCT-US95-15716-100	Sequence 100, App
c	5	8	100.0	20	1	US-08-420-244-10	Sequence 10, Appl
c	6	8	100.0	20	1	US-08-242-403A-4	Sequence 4, Appli
c	7	8	100.0	20	1	US-08-774-128-4	Sequence 4, Appli
	8	8	100.0	20	3	US-08-814-052-49	Sequence 49, Appl
c	9	8	100.0	20	3	US-08-754-490-15	Sequence 15, Appl
c	10	8	100.0	20	3	US-08-564-995-4	Sequence 4, Appli
	11	8	100.0	20	3	US-08-881-037-95	Sequence 95, Appl
	12	8	100.0	20	3	US-08-881-037-103	Sequence 103, App
c	13	8	100.0	20	3	US-08-922-505A-15	Sequence 15, Appl
c	14	8	100.0	20	4	US-09-260-952A-15	Sequence 15, Appl
c	15	8	100.0	20	4	US-09-253-341-15	Sequence 15, Appl
c	16	8	100.0	20	4	US-09-489-869-78	Sequence 78, Appl
c	17	8	100.0	20	4	US-09-253-331A-15	Sequence 15, Appl
c	18	8	100.0	20	5	PCT-US95-05602-4	Sequence 4, Appli
c	19	8	100.0	20	5	PCT-US95-05816-4	Sequence 4, Appli
	20	8	100.0	21	3	US-08-881-037-96	Sequence 96, Appl
	21	8	100.0	21	3	US-08-881-037-104	Sequence 104, App
	22	8	100.0	23	1	US-08-242-403A-5	Sequence 5, Appli
	23	8	100.0	23	1	US-08-774-128-5	Sequence 5, Appli
	24	8	100.0	23	3	US-08-564-995-5	Sequence 5, Appli
c	25	8	100.0	23	4	US-08-852-001-3	Sequence 3, Appli
c	26	8	100.0	23	4	US-08-852-001-10	Sequence 10, Appl
	27	8	100.0	23	5	PCT-US95-05602-5	Sequence 5, Appli
	28	8	100.0	23	5	PCT-US95-05816-5	Sequence 5, Appli
	29	8	100.0	24	1	US-07-882-838E-18	Sequence 18, Appl
c	30	8	100.0	26	1	US-08-220-606B-14	Sequence 14, Appl
	31	8	100.0	26	2	US-08-976-703-17	Sequence 17, Appl
	32	8	100.0	26	3	US-09-023-082A-108	Sequence 108, App
	33	8	100.0	26	4	US-09-218-444-29	Sequence 29, Appl
c	34	8	100.0	27	4	US-09-199-149-27	Sequence 27, Appl
	35	8	100.0	31	1	US-08-081-070-12	Sequence 12, Appl
	36	8	100.0	31	1	US-08-171-389-612	Sequence 612, App
	37	8	100.0	31	1	US-07-996-783-12	Sequence 12, Appl
	38	8	100.0	31	1	US-08-484-499-12	Sequence 12, Appl
	39	8	100.0	31	1	US-08-123-936-612	Sequence 612, App
	40	8	100.0	31	1	US-08-475-221B-12	Sequence 12, Appl
	41	8	100.0	31	1	US-08-476-876-12	Sequence 12, Appl
	42	8	100.0	31	2	US-08-475-228A-612	Sequence 612, App
	43	8	100.0	31	3	US-08-482-080A-612	Sequence 612, App
	44	8	100.0	31	4	US-09-070-408-117	Sequence 117, App
	45	8	100.0	31	5	PCT-US93-12388-612	Sequence 612, App

ALIGNMENTS

```
RESULT   1
US-08-485-355B-18/c
; Sequence 18, Application US/08485355B
; Patent No. 6177075
;   GENERAL INFORMATION:
;        APPLICANT: Christian, P. D., Gordon, K. H.J., Hanzlik, T. N.
;        TITLE OF INVENTION: Insect Viruses and Their Uses in
;                            Protecting Plants
;        NUMBER OF SEQUENCES: 57
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Flehr Hohbach Test Albritton & Herbert LLP
;             STREET: Four Embarcadero Center, Suite 3400
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: United States
;             ZIP: 94111-4187
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/485,355B
;             FILING DATE: 07-Jun-1995
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US 08/440,522
;             FILING DATE: 12-MAY-1995
;             APPLICATION NUMBER: US 08/089,372
;             FILING DATE: 08-JUL-1993
;             APPLICATION NUMBER: AU PL4081/92
;             FILING DATE: 14-AUG-1992
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Trecartin, Richard F.
;             REGISTRATION NUMBER: 31,801
;             REFERENCE/DOCKET NUMBER: A-58631-2/RFT/DSS
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 781-1989
;             TELEFAX: (415) 398-3249
;             TELEX: 910 277299
;   INFORMATION FOR SEQ ID NO: 18:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 17 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: unknown
;             TOPOLOGY: unknown
;        MOLECULE TYPE: DNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 18:
;
US-08-485-355B-18
```

```
  Query Match           100.0%;  Score 8;  DB 4;  Length 17;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      12 GACGTTCG 5


RESULT   2
US-09-025-769B-8/c
; Sequence 8, Application US/09025769B
; Patent No. 6300064
;  GENERAL INFORMATION:
;    APPLICANT:  Knappik, Achim
;    APPLICANT:  Pack, Peter
;    APPLICANT:  Ilag, Vic
;    APPLICANT:  Ge, Liming
;    APPLICANT:  Moroney, Simon
;    APPLICANT:  Plueckthun, Andreas
;    TITLE OF INVENTION:  Protein/(Poly)peptide libraries
;    NUMBER OF SEQUENCES:  373
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  James F. Haley, Jr., Esq. c/o Fish & Neave
;      STREET:  1251 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10021
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/025,769B
;      FILING DATE:  18-FEB-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  EP 95 11 3021.0
;      FILING DATE:  18-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  James F. Haley, Jr., Esq.
;      REGISTRATION NUMBER:  27,794
;      REFERENCE/DOCKET NUMBER:  MORPHO/5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212)596-9000
;      TELEFAX:  (212)596-9090
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  17 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "synthetic oligonucleotide"
US-09-025-769B-8


  Query Match           100.0%;  Score 8;  DB 4;  Length 17;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      16 GACGTTCG 9


RESULT   3
US-08-348-548-100/c
; Sequence 100, Application US/08348548
; Patent No. 6258529
;  GENERAL INFORMATION:
;    APPLICANT:  Berdoz, Jose
;    APPLICANT:  Kraehenbuhl, Jean Pierre
;    TITLE OF INVENTION:  PCR AMPLIFICATION OF REARRANGED GENOMIC
;    TITLE OF INVENTION:  VARIABLE REGIONS OF IMMUNOGLOBULIN GENES
;    NUMBER OF SEQUENCES:  108
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street, Suite 3100
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30B
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/348,548
;      FILING DATE:  01-DEC-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clark, Paul T.
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  06132/009001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-5070
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  100:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  19 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-08-348-548-100


  Query Match           100.0%;  Score 8;  DB 4;  Length 19;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      18 GACGTTCG 11


RESULT   4
PCT-US95-15716-100/c
; Sequence 100, Application PC/TUS9515716
;  GENERAL INFORMATION:
;    APPLICANT:  Berdoz, Jose
;    APPLICANT:  Kraehenbuhl, Jean Pierre
;    TITLE OF INVENTION:  PCR AMPLIFICATION OF REARRANGED GENOMIC
;    TITLE OF INVENTION:  VARIABLE REGIONS OF IMMUNOGLOBULIN GENES
;    NUMBER OF SEQUENCES:  108
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street, Suite 3100
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30B
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/15716
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/348,548
```

```
;      FILING DATE:  01-DEC-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clark, Paul T.
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  06132/009001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-5070
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  100:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  19 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
PCT-US95-15716-100


  Query Match             100.0%;  Score 8;  DB 5;  Length 19;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      18 GACGTTCG 11


RESULT   5
US-08-420-244-10/c
; Sequence 10, Application US/08420244
; Patent No. 5627195
;  GENERAL INFORMATION:
;    APPLICANT:  Hu, Shixing
;    TITLE OF INVENTION:  TREATMENT FOR OCULAR INFLAMMATION
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street, Suite 3100
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/420,244
;      FILING DATE:  07-APR-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Tsao, Y. Rocky
;      REGISTRATION NUMBER:  34,053
;      REFERENCE/DOCKET NUMBER:  00633/021001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-08-420-244-10


  Query Match             100.0%;  Score 8;  DB 1;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      16 GACGTTCG 9


RESULT   6
US-08-242-403A-4/c
; Sequence 4, Application US/08242403A
; Patent No. 5631130
;  GENERAL INFORMATION:
;    APPLICANT:  Leckie, G. W.
;    APPLICANT:  Davis, A. H.
;    APPLICANT:  Semple-Facey, I. E.
;    APPLICANT:  Manlove, M. T.
;    APPLICANT:  Solomon, N. A.
;    TITLE OF INVENTION:  Materials and Methods for the Detection of
;    TITLE OF INVENTION:  Mycobacteria tuberculosis
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Abbott Laboratories
;      STREET:  One Abbott Park Road
;      CITY:  Abbott Park
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60064-3500
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PCDOS/MSDOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/242,403A
;      FILING DATE:  May 13, 1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Thomas D. Brainard
;      REGISTRATION NUMBER:  32,459
;      REFERENCE/DOCKET NUMBER:  5370.US.01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  708/937-4884
;      TELEFAX:  708/938-2623
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  synthetic DNA
US-08-242-403A-4


  Query Match             100.0%;  Score 8;  DB 1;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      18 GACGTTCG 11


RESULT   7
US-08-774-128-4/c
; Sequence 4, Application US/08774128
; Patent No. 5786149
;  GENERAL INFORMATION:
;    APPLICANT:  Leckie, G. W.
;    APPLICANT:  Davis, A. H.
;    APPLICANT:  Semple-Facey, I. E.
;    APPLICANT:  Manlove, M. T.
;    APPLICANT:  Solomon, N. A.
;    TITLE OF INVENTION:  Materials and Methods for the Detection of
```

```
;    TITLE OF INVENTION:  Mycobacteria tuberculosis
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Abbott Laboratories
;      STREET:  One Abbott Park Road
;      CITY:  Abbott Park
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60064-3500
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PCDOS/MSDOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/774,128
;      FILING DATE:  23-DEC-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/242,403
;      FILING DATE:  May 13, 1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Thomas D. Brainard
;      REGISTRATION NUMBER:  32,459
;      REFERENCE/DOCKET NUMBER:  5370.US.01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  708/937-4884
;      TELEFAX:  708/938-2623
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  synthetic DNA
US-08-774-128-4


  Query Match             100.0%;  Score 8;  DB 1;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
            ||||||||
Db      18 GACGTTCG 11


RESULT   8
US-08-814-052-49
; Sequence 49, Application US/08814052
; Patent No. 6015783
;  GENERAL INFORMATION:
;    APPLICANT:  von der Osten, Claus
;    APPLICANT:  Cherry, Joel R.
;    APPLICANT:  Bjornvad, Mads E.
;    APPLICANT:  Vind, Jesper
;    APPLICANT:  Rasmussen, Michael Dolberg
;    TITLE OF INVENTION:  PROCESS FOR REMOVAL OR BLEACHING OF SOILING
;    TITLE OF INVENTION:  OR STAINS FROM CELLULOSIC FABRIC
;    NUMBER OF SEQUENCES:  55
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6015783o No. 6015783disk of No. 6015783th America, Inc.
;      STREET:  405 Lexington Avenue, Suite 6400
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10174-6401
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/814,052
;      FILING DATE:  06-MAR-1997
;      CLASSIFICATION:  510
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lambiris, Elias J
;      REGISTRATION NUMBER:  33,728
;      REFERENCE/DOCKET NUMBER:  4684.204-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-867-0123
;      TELEFAX:  212-878-9655
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  49:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-814-052-49


  Query Match             100.0%;  Score 8;  DB 3;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
            ||||||||
Db       9 GACGTTCG 16


RESULT   9
US-08-754-490-15/c
; Sequence 15, Application US/08754490
; Patent No. 6017534
;  GENERAL INFORMATION:
;    APPLICANT:  Malvar, Thomas
;    APPLICANT:  Gilmer, Amy Jelen
;    TITLE OF INVENTION:  HYBRID BACILLUS THURINGIENSIS
;    TITLE OF INVENTION:  DELTA-ENDOTOXINS WITH NOVEL BROAD SPECTRUM
;    TITLE OF INVENTION:  INSECTICIDAL ACTIVITY
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  United States of America
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/754,490
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kitchell, Barbara S.
;      REGISTRATION NUMBER:  33,928
;      REFERENCE/DOCKET NUMBER:  MOBT:009
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (512) 474-7577
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-754-490-15
```

```
  Query Match             100.0%;  Score 8;  DB 3;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      17 GACGTTCG 10


RESULT  10
US-08-564-995-4/c
; Sequence 4, Application US/08564995
; Patent No. 6071480
;  GENERAL INFORMATION:
;    APPLICANT:  Halaka, F.
;    TITLE OF INVENTION:  METHOD FOR GENERATING A STANDING SONIC WAVE, METHODS OF SONIC
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Abbott Laboratories
;      STREET:  100 Abbott Park Road
;      CITY:  Abbott Park
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60064-3500
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Macintosh
;      OPERATING SYSTEM:  System 7.0.1
;      SOFTWARE:  MS Word text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/564,995
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul D. Yasger
;      REGISTRATION NUMBER:  37,477
;      REFERENCE/DOCKET NUMBER:  5637.US.P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  708/938-3508
;      TELEFAX:  708/938-2623
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  synthetic DNA
US-08-564-995-4


  Query Match             100.0%;  Score 8;  DB 3;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      18 GACGTTCG 11


RESULT  11
US-08-881-037-95
; Sequence 95, Application US/08881037
; Patent No. 6080588
;  GENERAL INFORMATION:
;    APPLICANT:  Glick, Gary D.
;    APPLICANT:  Swanson, Patrick C.
;    TITLE OF INVENTION:  DNA BINDING ANTIBODIES
;    NUMBER OF SEQUENCES:  113
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morrison & Foerster
;      STREET:  755 Page Mill Road
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304-1018
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/881,037
;      FILING DATE:  23-JUN-1997
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/443,540
;      FILING DATE:  18-MAY-1995
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Konski, Antoinette F.
;      REGISTRATION NUMBER:  34,202
;      REFERENCE/DOCKET NUMBER:  203442110710
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 813-5600
;      TELEFAX:  (650) 494-0792
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  95:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  8..20
;      OTHER INFORMATION:  /note= "Portion of the germline
;      OTHER INFORMATION:  gene incorporated into the CDR3 construct"
US-08-881-037-95


  Query Match             100.0%;  Score 8;  DB 3;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      11 GACGTTCG 18


RESULT  12
US-08-881-037-103
; Sequence 103, Application US/08881037
; Patent No. 6080588
;  GENERAL INFORMATION:
;    APPLICANT:  Glick, Gary D.
;    APPLICANT:  Swanson, Patrick C.
;    TITLE OF INVENTION:  DNA BINDING ANTIBODIES
;    NUMBER OF SEQUENCES:  113
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morrison & Foerster
;      STREET:  755 Page Mill Road
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304-1018
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/881,037
;      FILING DATE:  23-JUN-1997
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
```

```
;       APPLICATION NUMBER:  US 08/443,540
;       FILING DATE:  18-MAY-1995
;       CLASSIFICATION:  530
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Konski, Antoinette F.
;       REGISTRATION NUMBER:  34,202
;       REFERENCE/DOCKET NUMBER:  203442110710
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (650) 813-5600
;       TELEFAX:  (650) 494-0792
;       TELEX:
;   INFORMATION FOR SEQ ID NO:  103:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  20 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  8..20
;       OTHER INFORMATION:  /note= "Portion of the germline
;       OTHER INFORMATION:  gene incorporated into the CDR3 construct"
US-08-881-037-103


  Query Match           100.0%;  Score 8;  DB 3;  Length 20;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      11 GACGTTCG 18


RESULT  13
US-08-922-505A-15/c
; Sequence 15, Application US/08922505A
; Patent No. 6110464
;  GENERAL INFORMATION:
;    APPLICANT:  Malvar, Thomas
;    APPLICANT:  Gilmer, Amy Jelen
;    TITLE OF INVENTION:  BROAD-SPECTRUM (-ENDOTOXINS
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/922,505A
;      FILING DATE:  03-SEP-1997
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kitchell, Barbara S.
;      REGISTRATION NUMBER:  33,928
;      REFERENCE/DOCKET NUMBER:  MECO:163
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512)418-3000
;      TELEFAX:  (512)474-7577
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-922-505A-15


  Query Match           100.0%;  Score 8;  DB 3;  Length 20;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      17 GACGTTCG 10


RESULT  14
US-09-260-952A-15/c
; Sequence 15, Application US/09260952A
; Patent No. 6221649
; GENERAL INFORMATION:
;  APPLICANT:  Malvar, Thomas
;  APPLICANT:  Gilmer, Amy Jelen
;  TITLE OF INVENTION:  HYBRID BACILLUS THURINGIENSIS DELTA-ENDOTOXINS WITH
;  TITLE OF INVENTION:  NOVEL BROAD SPECTRUM INSECTICIDAL ACTIVITY
;  FILE REFERENCE:  MECO:217
;  CURRENT APPLICATION NUMBER: US/09/260,952A
;  CURRENT FILING DATE:  1999-03-02
;  NUMBER OF SEQ ID NOS:  30
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 15
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: SYNTHETIC
US-09-260-952A-15


  Query Match           100.0%;  Score 8;  DB 4;  Length 20;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      17 GACGTTCG 10


RESULT  15
US-09-253-341-15/c
; Sequence 15, Application US/09253341
; Patent No. 6242241
;   GENERAL INFORMATION:
;       APPLICANT: Malvar, Thomas
;                  Gilmer, Amy Jelen
;       TITLE OF INVENTION: BROAD-SPECTRUM (-ENDOTOXINS
;       NUMBER OF SEQUENCES: 35
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Arnold, White & Durkee
;            STREET: P.O. Box 4433
;            CITY: Houston
;            STATE: Texas
;            COUNTRY: USA
;            ZIP: 77210
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/253,341
;            FILING DATE: 19-Feb-1999
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/922,505
;            FILING DATE: 03-SEP-1997
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Kitchell, Barbara S.
;            REGISTRATION NUMBER: 33,928
;            REFERENCE/DOCKET NUMBER: MECO:163
```

```
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (512)418-3000
;             TELEFAX: (512)474-7577
;    INFORMATION FOR SEQ ID NO: 15:
;         SEQUENCE CHARACTERISTICS:
;             LENGTH: 20 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;         SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-09-253-341-15


  Query Match             100.0%;  Score 8;  DB 4;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      17 GACGTTCG 10


RESULT  16
US-09-489-869-78/c
; Sequence 78, Application US/09489869A
; Patent No. 6268151
; GENERAL INFORMATION:
;  APPLICANT: Susan Murray
;  APPLICANT: Lex M. Cowsert
;  APPLICANT: Jacqueline Wyatt
;  TITLE OF INVENTION: ANTISENSE MODULATION OF MACROPHAGE MIGRATION INHIBITORY FACTOR
;  FILE REFERENCE: RTS-0110
;  CURRENT APPLICATION NUMBER: US/09/489,869A
;  CURRENT FILING DATE: 2000-01-20
;  NUMBER OF SEQ ID NOS: 88
; SEQ ID NO 78
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Antisense Oligonucleotide
US-09-489-869-78


  Query Match             100.0%;  Score 8;  DB 4;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      13 GACGTTCG 6


RESULT  17
US-09-253-331A-15/c
; Sequence 15, Application US/09253331A
; Patent No. 6281016
; GENERAL INFORMATION:
;  APPLICANT:  Malvar, Thomas
;  APPLICANT:  Gilmer, Amy Jelene
;  TITLE OF INVENTION:   BROAD-SPECTRUM INSECT RESISTANT TRANSGENIC PLANTS
;  FILE REFERENCE:   MECO211
;  CURRENT APPLICATION NUMBER: US/09/253,331A
;  CURRENT FILING DATE:  2000-02-19
;  PRIOR APPLICATION NUMBER:   08/922,505
;  PRIOR FILING DATE:   1997-09-03
;  NUMBER OF SEQ ID NOS:   35
;  SOFTWARE:   PatentIn version 3.0
; SEQ ID NO 15
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-253-331A-15


  Query Match             100.0%;  Score 8;  DB 4;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      17 GACGTTCG 10


RESULT  18
PCT-US95-05602-4/c
; Sequence 4, Application PC/TUS9505602
;  GENERAL INFORMATION:
;    APPLICANT:  Leckie, G.W.
;    APPLICANT:  Davis A.H.
;    APPLICANT:  Semple-Facey, I.E.
;    APPLICANT:  Manlove, M.T.
;    APPLICANT:  Solomon, N.A.
;    TITLE OF INVENTION:  Materials and Methods for the Detection of
;    TITLE OF INVENTION:  Mycobacteria tuberculosis
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Abbott Laboratories
;      STREET:  One Abbott Park Road
;      CITY:  Abbott Park
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60064-3500
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC_DOS/MS_DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/05602
;      FILING DATE:  May 13, 1994
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Thomas D. Brainard
;      REGISTRATION NUMBER:  32,459
;      REFERENCE/DOCKET NUMBER:  5370.PC.01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  708/937-4884
;      TELEFAX:  708/938-2623
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  synthetic DNA
PCT-US95-05602-4


  Query Match             100.0%;  Score 8;  DB 5;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      18 GACGTTCG 11


RESULT  19
PCT-US95-05816-4/c
; Sequence 4, Application PC/TUS9505816
;  GENERAL INFORMATION:
;    APPLICANT:  Solomon, N.
```

```
;    APPLICANT:  Leckie, G.
;    APPLICANT:  Kratochvil, J.
;    APPLICANT:  O'Donnell, D.
;    TITLE OF INVENTION:  Materials and Methods for the Detection of
;    TITLE OF INVENTION:  Mycobacteria
;    NUMBER OF SEQUENCES:  75
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Abbott Laboratories
;      STREET:  One Abbott Park Road
;      CITY:  Abbott Park
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60064-3500
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC_DOS/MS_DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/05816
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Thomas D. Brainard
;      REGISTRATION NUMBER:  32,459
;      REFERENCE/DOCKET NUMBER:  5371.PC.01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  708/937-4884
;      TELEFAX:  708/938-2623
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  synthetic DNA
PCT-US95-05816-4


  Query Match             100.0%;  Score 8;  DB 5;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      18 GACGTTCG 11


RESULT  20
US-08-881-037-96
; Sequence 96, Application US/08881037
; Patent No. 6080588
;  GENERAL INFORMATION:
;    APPLICANT:  Glick, Gary D.
;    APPLICANT:  Swanson, Patrick C.
;    TITLE OF INVENTION:  DNA BINDING ANTIBODIES
;    NUMBER OF SEQUENCES:  113
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morrison & Foerster
;      STREET:  755 Page Mill Road
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304-1018
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/881,037
;      FILING DATE:  23-JUN-1997
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/443,540
;      FILING DATE:  18-MAY-1995
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Konski, Antoinette F.
;      REGISTRATION NUMBER:  34,202
;      REFERENCE/DOCKET NUMBER:  203442110710
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 813-5600
;      TELEFAX:  (650) 494-0792
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  96:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..21
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  group(9, 10)
;      OTHER INFORMATION:  /note= "Positions that have mutated
;      OTHER INFORMATION:  away from the putative germline gene"
US-08-881-037-96


  Query Match             100.0%;  Score 8;  DB 3;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      12 GACGTTCG 19


RESULT  21
US-08-881-037-104
; Sequence 104, Application US/08881037
; Patent No. 6080588
;  GENERAL INFORMATION:
;    APPLICANT:  Glick, Gary D.
;    APPLICANT:  Swanson, Patrick C.
;    TITLE OF INVENTION:  DNA BINDING ANTIBODIES
;    NUMBER OF SEQUENCES:  113
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morrison & Foerster
;      STREET:  755 Page Mill Road
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304-1018
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/881,037
;      FILING DATE:  23-JUN-1997
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/443,540
;      FILING DATE:  18-MAY-1995
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Konski, Antoinette F.
;      REGISTRATION NUMBER:  34,202
;      REFERENCE/DOCKET NUMBER:  203442110710
;    TELECOMMUNICATION INFORMATION:
```

```
;       TELEPHONE:  (650) 813-5600
;       TELEFAX:  (650) 494-0792
;       TELEX:
;   INFORMATION FOR SEQ ID NO:  104:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  21 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  1..21
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  group(9, 11)
;       OTHER INFORMATION:  /note= "Positions that have mutated
;       OTHER INFORMATION:  away from the putative germline gene"
US-08-881-037-104


  Query Match             100.0%;  Score 8;  DB 3;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      12 GACGTTCG 19


RESULT  22
US-08-242-403A-5
; Sequence 5, Application US/08242403A
; Patent No. 5631130
;  GENERAL INFORMATION:
;    APPLICANT:  Leckie, G. W.
;    APPLICANT:  Davis, A. H.
;    APPLICANT:  Semple-Facey, I. E.
;    APPLICANT:  Manlove, M. T.
;    APPLICANT:  Solomon, N. A.
;    TITLE OF INVENTION:  Materials and Methods for the Detection of
;    TITLE OF INVENTION:  Mycobacteria tuberculosis
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Abbott Laboratories
;      STREET:  One Abbott Park Road
;      CITY:  Abbott Park
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60064-3500
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PCDOS/MSDOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/242,403A
;      FILING DATE:  May 13, 1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Thomas D. Brainard
;      REGISTRATION NUMBER:  32,459
;      REFERENCE/DOCKET NUMBER:  5370.US.01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  708/937-4884
;      TELEFAX:  708/938-2623
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  23 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  synthetic DNA
US-08-242-403A-5


  Query Match             100.0%;  Score 8;  DB 1;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db       3 GACGTTCG 10


RESULT  23
US-08-774-128-5
; Sequence 5, Application US/08774128
; Patent No. 5786149
;  GENERAL INFORMATION:
;    APPLICANT:  Leckie, G. W.
;    APPLICANT:  Davis, A. H.
;    APPLICANT:  Semple-Facey, I. E.
;    APPLICANT:  Manlove, M. T.
;    APPLICANT:  Solomon, N. A.
;    TITLE OF INVENTION:  Materials and Methods for the Detection of
;    TITLE OF INVENTION:  Mycobacteria tuberculosis
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Abbott Laboratories
;      STREET:  One Abbott Park Road
;      CITY:  Abbott Park
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60064-3500
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PCDOS/MSDOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/774,128
;      FILING DATE:  23-DEC-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/242,403
;      FILING DATE:  May 13, 1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Thomas D. Brainard
;      REGISTRATION NUMBER:  32,459
;      REFERENCE/DOCKET NUMBER:  5370.US.01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  708/937-4884
;      TELEFAX:  708/938-2623
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  23 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  synthetic DNA
US-08-774-128-5


  Query Match             100.0%;  Score 8;  DB 1;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db       3 GACGTTCG 10


RESULT  24
US-08-564-995-5
```

```
; Sequence 5, Application US/08564995
; Patent No. 6071480
;  GENERAL INFORMATION:
;    APPLICANT:  Halaka, F.
;    TITLE OF INVENTION:  METHOD FOR GENERATING A STANDING SONIC WAVE, METHODS OF SONIC
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Abbott Laboratories
;      STREET:  100 Abbott Park Road
;      CITY:  Abbott Park
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60064-3500
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Macintosh
;      OPERATING SYSTEM:  System 7.0.1
;      SOFTWARE:  MS Word text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/564,995
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul D. Yasger
;      REGISTRATION NUMBER:  37,477
;      REFERENCE/DOCKET NUMBER:  5637.US.P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  708/938-3508
;      TELEFAX:  708/938-2623
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  23 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  synthetic DNA
US-08-564-995-5


  Query Match             100.0%;  Score 8;  DB 3;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db       3 GACGTTCG 10


RESULT  25
US-08-852-001-3/c
; Sequence 3, Application US/08852001
; Patent No. 6197556
;  GENERAL INFORMATION:
;    APPLICANT:  Ulanovsky, Levy
;    APPLICANT:  Mugasimangalam, Raja C.
;    TITLE OF INVENTION:  NUCLEIC ACID AMPLIFICATION USING MODULAR
;    TITLE OF INVENTION:  BRANCHED PRIMERS
;    NUMBER OF SEQUENCES:  40
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BRINKS, HOFER, GILSON & LIONE
;      STREET:  NBC Tower - Suite 3600, 455 N. Cityfront
;      STREET:  Plaza Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60611-5599
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/852,001
;      FILING DATE:  06-MAY-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Martin, Alice O.
;      REGISTRATION NUMBER:  35,601
;      REFERENCE/DOCKET NUMBER:  6837/7
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-321-4200
;      TELEFAX:  312-321-4299
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  23 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "primer"
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1
;      OTHER INFORMATION:  /product= "N = inosine"
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  8
;      OTHER INFORMATION:  /product= "N = inosine"
US-08-852-001-3


  Query Match             100.0%;  Score 8;  DB 4;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      16 GACGTTCG 9


RESULT  26
US-08-852-001-10/c
; Sequence 10, Application US/08852001
; Patent No. 6197556
;  GENERAL INFORMATION:
;    APPLICANT:  Ulanovsky, Levy
;    APPLICANT:  Mugasimangalam, Raja C.
;    TITLE OF INVENTION:  NUCLEIC ACID AMPLIFICATION USING MODULAR
;    TITLE OF INVENTION:  BRANCHED PRIMERS
;    NUMBER OF SEQUENCES:  40
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BRINKS, HOFER, GILSON & LIONE
;      STREET:  NBC Tower - Suite 3600, 455 N. Cityfront
;      STREET:  Plaza Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60611-5599
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/852,001
;      FILING DATE:  06-MAY-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Martin, Alice O.
;      REGISTRATION NUMBER:  35,601
;      REFERENCE/DOCKET NUMBER:  6837/7
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-321-4200
;      TELEFAX:  312-321-4299
;  INFORMATION FOR SEQ ID NO:  10:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  23 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "primer"
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  8
;      OTHER INFORMATION:  /product= "N = inosine"
US-08-852-001-10
```

```
  Query Match             100.0%;  Score 8;  DB 4;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      16 GACGTTCG 9
```

```
RESULT  27
PCT-US95-05602-5
; Sequence 5, Application PC/TUS9505602
;  GENERAL INFORMATION:
;    APPLICANT:  Leckie, G.W.
;    APPLICANT:  Davis A.H.
;    APPLICANT:  Semple-Facey, I.E.
;    APPLICANT:  Manlove, M.T.
;    APPLICANT:  Solomon, N.A.
;    TITLE OF INVENTION:  Materials and Methods for the Detection of
;    TITLE OF INVENTION:  Mycobacteria tuberculosis
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Abbott Laboratories
;      STREET:  One Abbott Park Road
;      CITY:  Abbott Park
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60064-3500
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC_DOS/MS_DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/05602
;      FILING DATE:  May 13, 1994
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Thomas D. Brainard
;      REGISTRATION NUMBER:  32,459
;      REFERENCE/DOCKET NUMBER:  5370.PC.01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  708/937-4884
;      TELEFAX:  708/938-2623
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  23 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  synthetic DNA
PCT-US95-05602-5
```

```
  Query Match             100.0%;  Score 8;  DB 5;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db       3 GACGTTCG 10
```

```
RESULT  28
PCT-US95-05816-5
; Sequence 5, Application PC/TUS9505816
;  GENERAL INFORMATION:
;    APPLICANT:  Solomon, N.
;    APPLICANT:  Leckie, G.
;    APPLICANT:  Kratochvil, J.
;    APPLICANT:  O'Donnell, D.
;    TITLE OF INVENTION:  Materials and Methods for the Detection of
;    TITLE OF INVENTION:  Mycobacteria
;    NUMBER OF SEQUENCES:  75
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Abbott Laboratories
;      STREET:  One Abbott Park Road
;      CITY:  Abbott Park
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60064-3500
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC_DOS/MS_DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/05816
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Thomas D. Brainard
;      REGISTRATION NUMBER:  32,459
;      REFERENCE/DOCKET NUMBER:  5371.PC.01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  708/937-4884
;      TELEFAX:  708/938-2623
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  23 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  synthetic DNA
PCT-US95-05816-5
```

```
  Query Match             100.0%;  Score 8;  DB 5;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db       3 GACGTTCG 10
```

```
RESULT  29
US-07-882-838E-18
; Sequence 18, Application US/07882838E
; Patent No. 5616461
;  GENERAL INFORMATION:
;    APPLICANT:  Priscilla A. Schaffer
;    APPLICANT:  Christine E. Dabrowski Amaral
;    TITLE OF INVENTION:  METHODS AND COMPOSITIONS FOR
;    TITLE OF INVENTION:  TREATMENT OF VIRUS INFECTIONS
;    NUMBER OF SEQUENCES:  49
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn
;      STREET:  One Liberty Place
;      CITY:  Philadelphia
```

```
;       STATE:  Pennsylvania
;       COUNTRY:  U.S.A.
;       ZIP:  19103
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;       COMPUTER:  IBM PS/2 Model 50Z or 55SX
;       OPERATING SYSTEM:  IBM P.C. DOS (Version 3.30)
;       SOFTWARE:  WordPerfect (Version 5.1)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/882,838E
;       FILING DATE:  May 14, 1992
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Kathryn Leary
;       REGISTRATION NUMBER:  36,317
;       REFERENCE/DOCKET NUMBER:  DFCI-0001
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (215) 568-3100
;       TELEFAX:  (215) 568-3439
;       TELEX:
;   INFORMATION FOR SEQ ID NO:  18:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  24
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-07-882-838E-18


  Query Match           100.0%;  Score 8;  DB 1;  Length 24;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db       3 GACGTTCG 10


RESULT  30
US-08-220-606B-14/c
; Sequence 14, Application US/08220606B
; Patent No. 5641661
;  GENERAL INFORMATION:
;    APPLICANT:  Kumagai, Monto H.
;    APPLICANT:  Genadi, Sverlow J.
;    TITLE OF INVENTION:  Pichia Pastoris Alcohol Oxidase ZZA1 and
;    TITLE OF INVENTION:  ZZA2 Regulatory Regions for Heterologous Gene Expression
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/220,606B
;      FILING DATE:  25-MAR-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Halluin, Albert P.
;      REGISTRATION NUMBER:  25,227
;      REFERENCE/DOCKET NUMBER:  8129-065
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-854-3660
;      TELEFAX:  415-854-3694
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  26 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  DNA
US-08-220-606B-14


  Query Match           100.0%;  Score 8;  DB 1;  Length 26;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      22 GACGTTCG 15


RESULT  31
US-08-976-703-17
; Sequence 17, Application US/08976703
; Patent No. 5945288
;  GENERAL INFORMATION:
;    APPLICANT:  CHANG, ZHIYUH
;    APPLICANT:  MORGAN, RICHARD D.
;    TITLE OF INVENTION:  METHOD FOR CLONING AND
;    TITLE OF INVENTION:  PRODUCING THE PmeI RESTRICTION ENDONUCLEASE
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  New England Biolabs, Inc.
;      STREET:  32 Tozer Road
;      CITY:  Beverly
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  01915
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/976,703
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams, Gregory D
;      REGISTRATION NUMBER:  30901
;      REFERENCE/DOCKET NUMBER:  NEB-132
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  978-927-5054
;      TELEFAX:  978-927-1705
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  26 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
US-08-976-703-17


  Query Match           100.0%;  Score 8;  DB 2;  Length 26;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
```

```
          ||||||||
Db      12 GACGTTCG 19


RESULT  32
US-09-023-082A-108
; Sequence 108, Application US/09023082A
; Patent No. 6077692
;  GENERAL INFORMATION:
;    APPLICANT:  RUBEN, STEVEN M.
;    APPLICANT:  JIMENEZ, PABLO
;    APPLICANT:  DUAN, D. ROXANNE
;    APPLICANT:  RAMPY, MARK A.
;    APPLICANT:  MENDRICK, DONNA
;    APPLICANT:  ZHANG, JUN
;    APPLICANT:  NI, JIAN
;    APPLICANT:  MOORE, PAUL A.
;    APPLICANT:  COLEMAN, TIMOTHY A.
;    APPLICANT:  GRUBER, JOACHIM R.
;    APPLICANT:  DILLON, PATRICK J.
;    APPLICANT:  GENTZ, REINER L.
;    TITLE OF INVENTION:  KERATINOCYTE GROWTH FACTOR-2
;    NUMBER OF SEQUENCES:  148
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
;      STREET:  1100 NEW YORK AVE, NW, SUITE 600
;      CITY:  WASHINGTON
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/023,082A
;      FILING DATE:  13-FEB-1998
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/01790
;      FILING DATE:  14-FEB-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/461,195
;      FILING DATE:  05-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/023,852
;      FILING DATE:  13-AUG-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/039,045
;      FILING DATE:  28-FEB-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/862,432
;      FILING DATE:  23-MAY-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/910,875
;      FILING DATE:  13-AUG-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/055,561
;      FILING DATE:  13-AUG-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  STEFFFE, ERIC K.
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1488.0360008/EKS
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-371-2600
;      TELEFAX:  202-371-2540
;  INFORMATION FOR SEQ ID NO:  108:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  26 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-023-082A-108


  Query Match             100.0%;  Score 8;  DB 3;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      15 GACGTTCG 22


RESULT  33
US-09-218-444-29
; Sequence 29, Application US/09218444
; Patent No. 6238888
; GENERAL INFORMATION:
;  APPLICANT: Gentz, Reiner L.
;  APPLICANT: Chopra, Arvind
;  APPLICANT: Kaushal, Parveen
;  APPLICANT: Spitznagel, Thomas
;  APPLICANT: Unsworth, Edward
;  APPLICANT: Khan, Fazal
;  TITLE OF INVENTION: Keratinocyte Growth Factor-2 Formulations
;  FILE REFERENCE: 1488.1030001
;  CURRENT APPLICATION NUMBER: US/09/218,444
;  CURRENT FILING DATE: 1998-12-22
;  EARLIER APPLICATION NUMBER: US 60/068,493
;  EARLIER FILING DATE: 1997-12-22
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
;   LENGTH: 26
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-218-444-29


  Query Match             100.0%;  Score 8;  DB 4;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      15 gacgttcg 22


RESULT  34
US-09-199-149-27/c
; Sequence 27, Application US/09199149
; Patent No. 6160099
; GENERAL INFORMATION:
;  APPLICANT: Jonak, Zdenka L.
;  APPLICANT: Taylor, Alexander H.
;  APPLICANT: Trulli Jr., Stephen H.
;  APPLICANT: Johanson, Kyung O.
;  TITLE OF INVENTION: Humanized Monoclonal Antibodies
;  FILE REFERENCE: P50860
;  CURRENT APPLICATION NUMBER: US/09/199,149
;  CURRENT FILING DATE: 1998-11-24
;  NUMBER OF SEQ ID NOS: 37
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
;   LENGTH: 27
;   TYPE: DNA
;   ORGANISM: primer
US-09-199-149-27


  Query Match             100.0%;  Score 8;  DB 4;  Length 27;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy      1 GACGTTCG 8
          ||||||||
Db     22 GACGTTCG 15


RESULT  35
US-08-081-070-12
; Sequence 12, Application US/08081070
; Patent No. 5306619
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    TITLE OF INVENTION:  Screening Assay for the Detection of
;    TITLE OF INVENTION:  DNA-Binding Molecules
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dehlinger & Swiss
;      STREET:  P.O. Box 60850
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/081,070
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/723,618
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  4600-0085
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 323-8302
;      TELEFAX:  (415) 323-8306
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  31 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  oriECO2 TEST SEQ. / UL9 ASSAY SEQ.
US-08-081-070-12


  Query Match           100.0%;  Score 8;  DB 1;  Length 31;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db     10 GACGTTCG 17


RESULT  36
US-08-171-389-612
; Sequence 612, Application US/08171389
; Patent No. 5578444
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  641
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/171,389
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,070
;      FILING DATE:  22-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  4600-0175/G19P3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  612:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  31 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  oriECO2 TEST SEQ. / UL9 ASSAY SEQ.
US-08-171-389-612


  Query Match           100.0%;  Score 8;  DB 1;  Length 31;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db     10 GACGTTCG 17


RESULT  37
US-07-996-783-12
; Sequence 12, Application US/07996783
; Patent No. 5693463
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Fry, Kirk
;    TITLE OF INVENTION:  SEQUENCE-DIRECTED DNA-BINDING MOLECULES
;    TITLE OF INVENTION:  COMPOSITIONS AND METHODS
```

```
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Law Offices of Peter J. Dehlinger
;      STREET:  P.O. Box 60850
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/996,783
;      FILING DATE:  19921223
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  4600-0075.30
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  31 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  oriECO2 TEST SEQ. / UL9 ASSAY SEQ.
US-07-996-783-12


  Query Match             100.0%;  Score 8;  DB 1;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
            ||||||||
Db      10 GACGTTCG 17


RESULT  38
US-08-484-499-12
; Sequence 12, Application US/08484499
; Patent No. 5716780
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Fry, Kirk
;    TITLE OF INVENTION:  SEQUENCE-DIRECTED DNA-BINDING MOLECULES
;    TITLE OF INVENTION:  COMPOSITIONS AND METHODS
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Law Offices of Peter J. Dehlinger
;      STREET:  P.O. Box 60850
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,499
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  4600-0075.30
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  31 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  oriECO2 TEST SEQ. / UL9 ASSAY SEQ.
US-08-484-499-12


  Query Match             100.0%;  Score 8;  DB 1;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
            ||||||||
Db      10 GACGTTCG 17


RESULT  39
US-08-123-936-612
; Sequence 612, Application US/08123936
; Patent No. 5726014
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    TITLE OF INVENTION:  Screening Assay for the Detection of
;    TITLE OF INVENTION:  DNA-Binding Molecules
;    NUMBER OF SEQUENCES:  640
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/123,936
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  4600-0075.32/G19P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  612:
;    SEQUENCE CHARACTERISTICS:
```

```
;       LENGTH:  31 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     ORIGINAL SOURCE:
;       INDIVIDUAL ISOLATE:  oriECO2 TEST SEQ. / UL9 ASSAY SEQ.
US-08-123-936-612


  Query Match             100.0%;  Score 8;  DB 1;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      10 GACGTTCG 17


RESULT  40
US-08-475-221B-12
; Sequence 12, Application US/08475221B
; Patent No. 5738990
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A
;    APPLICANT:  Fry, Kirk E
;    APPLICANT:  Cantor, Charles R
;    APPLICANT:  Andrews, Beth M
;    TITLE OF INVENTION:  Sequence-Directed DNA-Binding Molecules
;    TITLE OF INVENTION:  Compositions and Methods
;    NUMBER OF SEQUENCES:  50
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dehlinger & Associates
;      STREET:  350 Cambridge Ave., Suite 250
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/475,221B
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stratford, Carol A
;      REGISTRATION NUMBER:  34,444
;      REFERENCE/DOCKET NUMBER:  4600-0075.34
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-324-0880
;      TELEFAX:  415-324-0960
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  31 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  oriECO2 TEST SEQ. / UL9 ASSAY SEQ.
US-08-475-221B-12


  Query Match             100.0%;  Score 8;  DB 1;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      10 GACGTTCG 17


RESULT  41
US-08-476-876-12
; Sequence 12, Application US/08476876
; Patent No. 5744131
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A
;    APPLICANT:  Fry, Kirk E
;    APPLICANT:  Cantor, Charles R
;    APPLICANT:  Andrews, Beth M
;    TITLE OF INVENTION:  Sequence-Directed DNA-Binding Molecules
;    TITLE OF INVENTION:  Compositions and Methods
;    NUMBER OF SEQUENCES:  50
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dehlinger & Associates
;      STREET:  350 Cambridge Ave., Suite 250
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/476,876
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stratford, Carol A
;      REGISTRATION NUMBER:  34,444
;      REFERENCE/DOCKET NUMBER:  4600-0075.33
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-324-0880
;      TELEFAX:  415-324-0960
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  31 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  oriECO2 TEST SEQ. / UL9 ASSAY SEQ.
US-08-476-876-12


  Query Match             100.0%;  Score 8;  DB 1;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      10 GACGTTCG 17


RESULT  42
US-08-475-228A-612
; Sequence 612, Application US/08475228A
```

```
; Patent No. 5869241
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  664
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/475,228A
;      FILING DATE:  06-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,070
;      FILING DATE:  22-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stratford, Carol A.
;      REGISTRATION NUMBER:  34,444
;      REFERENCE/DOCKET NUMBER:  4600-0175.21/G19P3D2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  612:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  31 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  oriECO2 TEST SEQ. / UL9 ASSAY SEQ.
US-08-475-228A-612


  Query Match           100.0%;  Score 8;  DB 2;  Length 31;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
            ||||||||
Db      10 GACGTTCG 17


RESULT  43
US-08-482-080A-612
; Sequence 612, Application US/08482080A
; Patent No. 6010849
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  664
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/482,080A
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/171,389
;      FILING DATE:  20-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,070
;      FILING DATE:  22-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brady, John F.
;      REGISTRATION NUMBER:  39,118
;      REFERENCE/DOCKET NUMBER:  4600-0175.20/G19P3D1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 324-0880
;      TELEFAX:  (650) 324-0960
;  INFORMATION FOR SEQ ID NO:  612:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  31 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  oriECO2 TEST SEQ. / UL9 ASSAY SEQ.
US-08-482-080A-612


  Query Match           100.0%;  Score 8;  DB 3;  Length 31;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
            ||||||||
Db      10 GACGTTCG 17


RESULT  44
US-09-070-408-117
; Sequence 117, Application US/09070408
; Patent No. 6180341
;  GENERAL INFORMATION:
;    APPLICANT:  Iverson, Brent L.
```

```
;    APPLICANT:  Georgiou, George
;    APPLICANT:  Burks, Elizabeth A.
;    TITLE OF INVENTION:  IN VITRO SCANNING SATURATION MUTAGENESIS
;    TITLE OF INVENTION:  OF PROTEINS
;    NUMBER OF SEQUENCES:  132
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/070,408
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/045,409
;      FILING DATE:  01-MAY-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMillian, Nabeela R.
;      REGISTRATION NUMBER:  P-43,363
;      REFERENCE/DOCKET NUMBER:  UTSB:593
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/418-3000
;      TELEFAX:  512/447-7577
;  INFORMATION FOR SEQ ID NO:  117:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  31 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-070-408-117


  Query Match             100.0%;  Score 8;  DB 4;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
            ||||||||
Db      24 GACGTTCG 31


RESULT  45
PCT-US93-12388-612
; Sequence 612, Application PC/TUS9312388
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  641
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/12388
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  4600-0175.41/G19PCT2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  612:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  31 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  oriECO2 TEST SEQ. / UL9 ASSAY SEQ.
PCT-US93-12388-612


  Query Match             100.0%;  Score 8;  DB 5;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
            ||||||||
Db      10 GACGTTCG 17


Search completed:  November 29, 2001, 14:48:18
Job time:  3591 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: November 29, 2001, 14:23:46 ; Search time 1878.42 Seconds
(without alignments)
45.765 Million cell updates/sec

Title: FRAG2
Perfect score: 8
Sequence: 1 GACGTTCG 8

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 260912

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_htc:*
10: gb_est1:*
11: gb_est2:*
12: gb_htc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	8	100.0	22	10	AA991491	AA991491 os91h12.s
	2	8	100.0	28	10	AI441029	AI441029 sa58e02.y
c	3	8	100.0	39	13	TA74E07P	AL457643 T. brucei
	4	8	100.0	40	11	BG177506	BG177506 602314157
c	5	8	100.0	41	13	CNS00BFE	AL057007 Drosophil
	6	8	100.0	45	11	BF538233	BF538233 602053710
	7	8	100.0	46	10	AA399336	AA399336 zt49c12.r
	8	8	100.0	49	10	AA519644	AA519644 TgESTzz42
	9	8	100.0	50	10	AU104010	AU104010 AU104010
	10	8	100.0	50	10	AU105830	AU105830 AU105830
	11	8	100.0	50	10	AU105831	AU105831 AU105831
c	12	8	100.0	58	10	AA574519	AA574519 vm29c07.r
	13	8	100.0	58	11	BF117267	BF117267 uz07c01.y
	14	8	100.0	59	11	BI175065	BI175065 OSTR007C1
c	15	8	100.0	60	13	CNS01356	AL102420 Drosophil
	16	8	100.0	63	10	AA486663	AA486663 ab16d10.r
	17	8	100.0	64	10	AI748274	AI748274 sb50d01.y
	18	8	100.0	64	10	AA458519	AA458519 zx96b04.r
	19	8	100.0	64	10	AA606766	AA606766 vm86e04.r
	20	8	100.0	64	11	H53706	H53706 yu38c10.r1
	21	8	100.0	65	13	TA303C12Q	AL489246 T. brucei
c	22	8	100.0	66	10	AW063368	AW063368 TN0743 KR
c	23	8	100.0	67	10	AA688966	AA688966 vs01a10.r
c	24	8	100.0	70	10	AI956572	AI956572 ul78h08.y
c	25	8	100.0	70	10	AI986743	AI986743 rs16h09.y
	26	8	100.0	71	10	AI365158	AI365158 qx97a08.x
c	27	8	100.0	72	10	AA243275	AA243275 zr26g11.r
c	28	8	100.0	73	10	AA967742	AA967742 uh04c05.r
c	29	8	100.0	73	10	AI180756	AI180756 ub91f11.r
c	30	8	100.0	73	10	AA120541	AA120541 mn12c03.r
c	31	8	100.0	76	10	AA863599	AA863599 vx05e04.r
c	32	8	100.0	78	10	AW717722	AW717722 i1b08nm.f
	33	8	100.0	86	10	AA284594	AA284594 zt22b11.r
	34	8	100.0	91	10	AA239711	AA239711 myl5e08.r
	35	8	100.0	92	10	AA634931	AA634931 ab28h02.r
c	36	8	100.0	94	10	AA661504	AA661504 nr18c12.s
c	37	8	100.0	94	10	AI957911	AI957911 fd08a06.x
c	38	8	100.0	95	10	AA722920	AA722920 zg81c01.s
	39	8	100.0	95	10	AW630585	AW630585 hh81g04.y
	40	8	100.0	95	10	AA426003	AA426003 zw17e07.r
	41	8	100.0	98	10	AI329158	AI329158 b1g10ne.r
	42	8	100.0	99	10	AA593996	AA593996 nn16f03.s
c	43	8	100.0	100	10	AI622446	AI622446 486055C03
	44	8	100.0	100	11	BG272807	BG272807 nah90g06.
c	45	8	100.0	100	13	TA101B10P	AL458854 T. brucei

ALIGNMENTS

RESULT 1
AA991491
LOCUS AA991491 22 bp mRNA EST 03-JUN-1998
DEFINITION os91h12.s1 NCI_CGAP_GC3 Homo sapiens cDNA clone IMAGE:1612775 3' similar to TR:O14597 O14597 NON-FUNCTIONAL FOLATE BINDING PROTEIN. ;, mRNA sequence.
ACCESSION AA991491
VERSION AA991491.1 GI:3177980
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 22)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .22
/organism="Homo sapiens"

/db_xref="taxon:9606"
/clone="IMAGE:1612775"
/clone_lib="NCI_CGAP_GC3"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from 3 pooled germ cell tumors, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is not normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 4 a 4 c 9 g 5 t
ORIGIN

Query Match 100.0%; Score 8; DB 10; Length 22;
Best Local Similarity 100.0%; Pred. No. 2.9e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGTTCG 8
||||||||
Db 7 GACGTTCG 14

RESULT 2
AI441029
LOCUS AI441029 28 bp mRNA EST 01-DEC-1999
DEFINITION sa58e02.y1 Gm-c1004 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-c1004-3507 5' similar to TR:Q41454 Q41454 HMG-COA REDUCTASE ;, mRNA sequence.
ACCESSION AI441029
VERSION AI441029.1 GI:4286315
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE 1 (bases 1 to 28)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomesystems.com web site: www.genomesystems.com
Possible reversed clone: similarity on wrong strand
Seq primer: -40RP from Gibco
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .28
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1004-3507"
/clone_lib="Gm-c1004"
/tissue_type="root"
/lab_host="XL10-Gold"
/note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2: XhoI; Root cDNA. The mRNA was isolated from entire roots of 8 day old 'Williams' seedlings which were propagated on paper towels with distilled water. Stratagene's cDNA Synthesis Kit (catalog #200401) was used to synthesize the cDNA. First- strand synthesis was performed with 5-methyl dCTP, hence the ligated cDNA is hemimethylated. Stratagene's first-strand synthesis primer was used [GAGAGAGAGAGAGAGAGAACTAGTCTCGAG(T)-18]. After second-strand synthesis, the cDNA ends were 'polished' with clone Pfu DNA polymerase, ligated to EcoRI adapters, and phosphorylated. The XhoI site within the first-strand synthesis primer was restricted by digestion with XhoI; all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500bp cutoff, using GibcoBRL Life Technologies' cDNA Size Fractionation column. The column eluent was then ligated into Stratagene's pBluescript II XR Predigested vector (pBluescript II SK(+) that had been digested with EcoRI and XhoI, and phosphorylated). Both the white and blue colonies appear to contain recombinant plasmids with cDNA inserts. Blue colonies 9n=15) have been sequenced, and possess putative cDNA inserts. This library was constructed by Dr. Paul Keim & Virginia H. Coryell, Department of Biology, Box5640, Northern Arizona University, Flagstaff, AZ 86011, Phone: 520-523-1078 (Dr. Paul Keim), 520-523-1372 (Virginia H. Coryell), Fax: 520-523-7500, email: paul.keim@nau.edu, virginia.coryell@nau.edu"

BASE COUNT 4 a 12 c 7 g 5 t
ORIGIN

Query Match 100.0%; Score 8; DB 10; Length 28;
Best Local Similarity 100.0%; Pred. No. 3e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGTTCG 8
||||||||
Db 2 GACGTTCG 9

RESULT 3
TA74E07P/c
LOCUS TA74E07P 39 bp DNA GSS 13-DEC-2000
DEFINITION T. brucei sheared genomic DNA clone 74e07, forward sequence, genomic survey sequence.
ACCESSION AL457643
VERSION AL457643.1 GI:11859606
KEYWORDS GSS.
SOURCE Trypanosoma brucei.
ORGANISM Trypanosoma brucei
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
REFERENCE 1 (bases 1 to 39)
AUTHORS Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R., Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L., Melville,S.E., Rajandream,M.A. and Barrell,B.G.
TITLE Direct Submission
JOURNAL Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and nh1@sanger.ac.uk
COMMENT Constructed at the Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from a cloned population of Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared to give a tight size distribution (
4 kb). The v + i method used for the library construction is described in detail in Smith, H. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999).
Email: nelsayed@tigr.org

```
          Details of T. brucei sequencing at the Sanger Centre are available
          at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES             Location/Qualifiers
     source          1. .39
                     /organism="Trypanosoma brucei"
                     /strain="TREU927"
                     /db_xref="taxon:5691"
                     /clone="74e07"
BASE COUNT        9 a      14 c       9 g       7 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 13;  Length 39;
  Best Local Similarity   100.0%;  Pred. No. 3.2e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      25 GACGTTCG 18


RESULT   4
BG177506
LOCUS       BG177506       40 bp    mRNA            EST       06-FEB-2001
DEFINITION  602314157F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4419759 5',
            mRNA sequence.
ACCESSION   BG177506
VERSION     BG177506.1  GI:12684209
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 40)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis Staudt, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10156  row: b  column: 16
            High quality sequence stop: 40.
FEATURES             Location/Qualifiers
     source          1. .40
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4419759"
                     /clone_lib="NIH_MGC_85"
                     /tissue_type="lymphoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.867 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT        8 a      10 c       13 g      9 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 11;  Length 40;
  Best Local Similarity   100.0%;  Pred. No. 3.3e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db       5 GACGTTCG 12


RESULT   5
CNS00BFE/c
LOCUS       CNS00BFE       41 bp    DNA             GSS       04-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR23C24 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL057007
VERSION     AL057007.1  GI:4937574
KEYWORDS    GSS.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 41)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .41
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone_lib="RPCI-98"
                     /clone="BACR23C24"
                     /note="end : TET3"
BASE COUNT       11 a       6 c       8 g       5 t      11 others
ORIGIN


  Query Match             100.0%;  Score 8;  DB 13;  Length 41;
  Best Local Similarity   100.0%;  Pred. No. 3.3e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      32 GACGTTCG 25


RESULT   6
BF538233
LOCUS       BF538233       45 bp    mRNA            EST       11-DEC-2000
DEFINITION  602053710F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4192827 5',
            mRNA sequence.
ACCESSION   BF538233
VERSION     BF538233.1  GI:11625601
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
```

```
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9524  row: c  column: 04
            High quality sequence stop: 45.
FEATURES              Location/Qualifiers
     source           1. .45
                      /organism="Mus musculus"
                      /strain="FVB/N"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:4192827"
                      /clone_lib="NCI_CGAP_SG2"
                      /lab_host="DH10B (T1 phage-resistant)"
                      /note="Organ: salivary gland; Vector: pCMV-SPORT6; Site_1:
                      NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo
                      dT. Average insert size 1.3 kb. Constructed by Life
                      Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT         6 a      10 c      22 g       7 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 11;  Length 45;
  Best Local Similarity   100.0%;  Pred. No. 3.3e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
            ||||||||
Db      13 GACGTTCG 20


RESULT   7
AA399336
LOCUS        AA399336        46 bp    mRNA            EST       08-AUG-1997
DEFINITION   zt49c12.r1 Soares ovary tumor NbHOT Homo sapiens cDNA clone
             IMAGE:725686 5' similar to gb:X72467 IG KAPPA CHAIN PRECURSOR V-II
             REGION (HUMAN);, mRNA sequence.
ACCESSION    AA399336
VERSION      AA399336.1  GI:2053073
KEYWORDS     EST.
SOURCE       human.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 46)
  AUTHORS    Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
             Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
             ,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
             ,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
             Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E.,
             Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
  TITLE      Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL    Genome Res. 6 (9), 807-828 (1996)
  MEDLINE    97044478
COMMENT      Contact: Wilson RK
             Washington University School of Medicine
             4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
             Tel: 314 286 1800
             Fax: 314 286 1810
             Email: est@watson.wustl.edu
             This clone is available royalty-free through LLNL ; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.
             Insert Length: 944    Std Error: 0.00
             Seq primer: -28ml3 rev2 ET from Amersham
             High quality sequence stop: 28.
FEATURES              Location/Qualifiers
     source           1. .46
                      /organism="Homo sapiens"
                      /db_xref="GDB:5937605"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:725686"
                      /clone_lib="Soares ovary tumor NbHOT"
                      /sex="Female"
                      /tissue_type="ovarian tumor"
                      /lab_host="DH10B (ampicillin resistant)"
                      /note="Organ: ovary; Vector: pT7T3D (Pharmacia) with a
                      modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                      strand cDNA was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCGGTTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was size selected, ligated to Eco RI
                      adapters (Pharmacia), digested with Not I and cloned into
                      the Not I and Eco RI sites of a modified pT7T3 vector
                      (Pharmacia). Library constructed by Bento Soares and
                      M.Fatima Bonaldo. "
BASE COUNT        14 a       9 c      16 g       7 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 46;
  Best Local Similarity   100.0%;  Pred. No. 3.3e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
            ||||||||
Db       4 GACGTTCG 11


RESULT   8
AA519644
LOCUS        AA519644        49 bp    mRNA            EST       22-MAY-2000
DEFINITION   TgESTzz42h05.s1 TgME49 invivo Bradyzoite cDNA size selected
             Toxoplasma gondii cDNA clone tgzz42h05.s1 3' similar to TR:G971750
             G971750 RIBOSOMAL PROTEIN S16 ;, mRNA sequence.
ACCESSION    AA519644
VERSION      AA519644.1  GI:2260048
KEYWORDS     EST.
SOURCE       Toxoplasma gondii.
  ORGANISM   Toxoplasma gondii
             Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
             Sarcocystidae; Toxoplasma.
REFERENCE    1  (bases 1 to 49)
  AUTHORS    Hehl,A., Manger,I., Marra,M., Parmley,S., Sibley,L.D., Hillier,L.,
             Allen,M., Bowles,L., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le
             ,N., Jost,S., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan
             ,F., Theising,B., Bowers,Y., Wylie,T., Ajioka,J.A., Aslett,M.A.,
             Wan,K.L., Wilson,R., Waterston,R. and Boothroyd J.C.
  TITLE      WashU-Stanford-PAMF-NIH Toxoplasma EST project
  JOURNAL    Unpublished (1997)
COMMENT      Contact: Marra M
             WashU-Merck EST Project
             Washington University School of Medicine
             4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
             Tel: 314 286 1800
             Fax: 314 286 1810
             Email: toxo@watson.wustl.edu
             Contact John Boothroyd (jboothr@leland.Stanford.EDU) for
             information on clone and library availability.
             Trace considered overall poor quality
             Possible reversed clone: similarity on wrong strand
             High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .49
                      /organism="Toxoplasma gondii"
                      /strain="ME49"
                      /db_xref="taxon:5811"
                      /clone="tgzz42h05.s1"
                      /clone_lib="TgME49 invivo Bradyzoite cDNA size selected"
                      /dev_stage="Bradyzoite"
                      /lab_host="DH10"
                      /note="Vector: Bluescript II SK-; Site_1: EcoRI; Site_2:
                      NotI; Mature bradyzoites were obtained from infected mouse
                      brains by percoll density centrifugation. The original
```

```
                 library was constructed by Steve Parmley, Palo Alto
                 Medical Foundation. cDNAs were synthesized by priming with
                 oligo d(T) and directionally cloned into the EcoRI/NotI
                 sites of lambda gt11. Warning: the library contains a
                 small percentage of host cDNAs derived from mouse cells.
                 Inserts from this cDNA library were excised with Not1 and
                 EcoR1, size selected in a range of 0.7 - 2.0 kb and
                 subcloned into Bluescript II SK- (Adrian Hehl, Ian Manger
                 and John Boothroyd, Stanford University)"
BASE COUNT       12 a     14 c     16 g      7 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 49;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
            ||||||||
Db      16 GACGTTCG 23


RESULT   9
AU104010
LOCUS       AU104010       50 bp    mRNA            EST       05-APR-2001
DEFINITION  AU104010 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HEP12302, mRNA sequence.
ACCESSION   AU104010
VERSION     AU104010.1  GI:13553531
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
            ,K., Suyama,A. and Sugano,S.
  TITLE     Fine Structural analysis of transcription start sites of human
            mRNAs using full-length enriched and 5'-end enriched cDNA libraries
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HEP12302"
                     /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT       9 a     15 c     16 g     10 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 50;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
            ||||||||
Db      21 GACGTTCG 28


RESULT  10
AU105830
LOCUS       AU105830       50 bp    mRNA            EST       05-APR-2001
DEFINITION  AU105830 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HSI05479, mRNA sequence.
ACCESSION   AU105830
VERSION     AU105830.1  GI:13555351
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
            ,K., Suyama,A. and Sugano,S.
  TITLE     Fine Structural analysis of transcription start sites of human
            mRNAs using full-length enriched and 5'-end enriched cDNA libraries
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HSI05479"
                     /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT       3 a     18 c     14 g     15 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 50;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
            ||||||||
Db      32 GACGTTCG 39


RESULT  11
AU105831
LOCUS       AU105831       50 bp    mRNA            EST       05-APR-2001
DEFINITION  AU105831 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HSI06453, mRNA sequence.
ACCESSION   AU105831
VERSION     AU105831.1  GI:13555352
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
            ,K., Suyama,A. and Sugano,S.
  TITLE     Fine Structural analysis of transcription start sites of human
            mRNAs using full-length enriched and 5'-end enriched cDNA libraries
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HSI06453"
```

```
                     /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT        3 a      18 c      14 g      15 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 50;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      32 GACGTTCG 39


RESULT  12
AA574519/c
LOCUS       AA574519       58 bp    mRNA            EST       02-SEP-1997
DEFINITION  vm29c07.r1 Knowles Solter mouse blastocyst B1 Mus musculus cDNA
            clone IMAGE:991596 5' similar to SW:RL2B_HUMAN P39024 60S RIBOSOMAL
            PROTEIN L23A;, mRNA sequence.
ACCESSION   AA574519
VERSION     AA574519.1  GI:2349145
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 58)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:563876
            Trace considered overall poor quality
            Putative full length read
            vector to vector length is 423
            Possible reversed clone: similarity on wrong strand
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .58
                     /organism="Mus musculus"
                     /strain="B6D2 F1/J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:991596"
                     /clone_lib="Knowles Solter mouse blastocyst B1"
                     /tissue_type="blastocyst"
                     /dev_stage="embryo (pre-implantation)"
                     /lab_host="DH10B"
                     /note="Organ: embryo; Vector: pSPORT; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally from mRNA prepared
                     from 800 blastocysts.  Primer: SalI(dT):
                     5'-CGGTCGACCGTCGACCGTTTTTTTTTTTTTT-3'.  cDNAs  were
                     cloned into the NotI/SalI sites of a pSPORT vector (Life
                     Technologies).  Two different size selections:  B1 (larger
                     inserts) and B3."
BASE COUNT       20 a      21 c      12 g       5 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 58;
  Best Local Similarity   100.0%;  Pred. No. 3.5e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      36 GACGTTCG 29


RESULT  13
BF117267
LOCUS       BF117267       58 bp    mRNA            EST       29-DEC-2000
DEFINITION  uz07c01.y1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:3668352 5',
            mRNA sequence.
ACCESSION   BF117267
VERSION     BF117267.1  GI:10986715
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 58)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            image.llnl.gov/image/html/iresources.shtml

            MGI:1429120
            Seq primer: -40RP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .58
                     /organism="Mus musculus"
                     /strain="C57/B6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3668352"
                     /clone_lib="NCI_CGAP_Mam5"
                     /tissue_type="tumor, gross tissue"
                     /dev_stage="7 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigators
                     providing samples: Lothar Hennighausen/Robin Humphreys,
                     NIH"
BASE COUNT       13 a      11 c      19 g      15 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 11;  Length 58;
  Best Local Similarity   100.0%;  Pred. No. 3.5e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      22 GACGTTCG 29


RESULT  14
BI175065
LOCUS       BI175065       59 bp    mRNA            EST       09-JUL-2001
DEFINITION  OSTR007C11_1 AD-wrmcDNA Caenorhabditis elegans cDNA similar to
            E03D2.1, mRNA sequence.
ACCESSION   BI175065
VERSION     BI175065.1  GI:14640868
KEYWORDS    EST.
SOURCE      Caenorhabditis elegans.
```

```
  ORGANISM  Caenorhabditis elegans
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
            ; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE   1  (bases 1 to 59)
  AUTHORS   Reboul,J., Vaglio,P., Tzellas,N., Thierry-Mieg,N., Moore,T.,
            Jackson,C., Shin-i,T., Kohara,Y., Thierry-Mieg,D., Thierry-Mieg,J.,
            Lee,H., Hitti,J., Doucette-Stamm,L., Hartley,J.L., Temple,G.F.,
            Brasch,M.A., Vandenhaute,J., Lamesch,P.E., Hill,D.E. and Vidal,M.
  TITLE     Open-reading-frame sequence tags (OSTs) support the existence of at
            least 17,300 genes in C. elegans
  JOURNAL   Nat. Genet. 27 (3), 332-336 (2001)
  MEDLINE   21135099
COMMENT     Contact: Reboul J, Vaglio P
            Marc Vidal Laboratory
            Dana Farber Cancer Institute
            44 Binney Street, Boston, MA 02115, USA
            Tel: 617 632 5180
            Fax: 617 632 2425
            Email: Jerome_Reboul@dfci.harvard.edu
            Sequence tag of Gateway entry clones. The primers used were
            designed on the predicted protein encoding ORF. C. elegans ORFeome
            cloning project : Contact jerome_reboul@dfci.harvard.edu or
            philippe_vaglio@dfci.harvard.edu
            POLYA=No.
FEATURES              Location/Qualifiers
     source           1. .59
                      /organism="Caenorhabditis elegans"
                      /strain="N2"
                      /db_xref="taxon:6239"
                      /clone_lib="AD-wrmcDNA"
                      /sex="Hermaphrodite and male"
                      /tissue_type="whole animal"
                      /dev_stage="mixed stage"
                      /note="The AD-wrmcDNA library was generated with poly(A)+
                      RNA isolated from both hermaphrodite and male N2 worms of
                      all larval stages, embryos, adults and dauers and the
                      subsequent generation of cDNAs by poly(A) priming. The
                      cDNAs were cloned into pPC86"
BASE COUNT       11 a     19 c     13 g     16 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 11;  Length 59;
  Best Local Similarity   100.0%;  Pred. No. 3.5e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      10 GACGTTCG 17


RESULT  15
CNS01356/c
LOCUS       CNS01356       60 bp    DNA             GSS       26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence T7 end of BAC
            BACN09E03 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL102420
VERSION     AL102420.1  GI:5614031
KEYWORDS    GSS.
SOURCE      fruit fly.
  ORGANISM  Plasmid Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBeloBAC11.
FEATURES              Location/Qualifiers
     source           1. .60
                      /organism="Drosophila melanogaster"
                      /plasmid="pBeloBAC11"
                      /db_xref="taxon:7227"
                      /clone_lib="DrosBAC"
                      /clone="BACN09E03"
                      /note="end : T7"
BASE COUNT       15 a     10 c      7 g     13 t     15 others
ORIGIN


  Query Match             100.0%;  Score 8;  DB 13;  Length 60;
  Best Local Similarity   100.0%;  Pred. No. 3.5e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      17 GACGTTCG 10


RESULT  16
AA486663
LOCUS       AA486663       63 bp    mRNA            EST       06-MAR-1998
DEFINITION  ab16d10.r1 Stratagene lung (#937210) Homo sapiens cDNA clone
            IMAGE:840979 5' similar to gb:X72467 IG KAPPA CHAIN PRECURSOR V-II
            REGION (HUMAN);, mRNA sequence.
ACCESSION   AA486663
VERSION     AA486663.1  GI:2216827
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 63)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 810   Std Error: 0.00
            Seq primer: -28m13 rev1 ET from Amersham.
FEATURES              Location/Qualifiers
     source           1. .63
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:840979"
                      /clone_lib="Stratagene lung (#937210)"
                      /sex="male"
                      /dev_stage="72 years"
                      /lab_host="SOLR cells (kanamycin resistant)"
                      /note="Organ: lung; Vector: pBluescript SK-; Site_1: EcoRI
                      ; Site_2: XhoI; Cloned unidirectionally.  Primer: Oligo
                      dT. normal lung.  Average insert size: 1.0 kb; Uni-ZAP XR
                      Vector; -5' adaptor sequence: 5' GAATTCGGCACGAG 3' -3'
                      adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT       20 a     14 c     17 g     12 t
ORIGIN
```

```
  Query Match             100.0%;  Score 8;  DB 10;  Length 63;
  Best Local Similarity   100.0%;  Pred. No. 3.5e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      21 GACGTTCG 28


RESULT  17
AI748274
LOCUS       AI748274       64 bp    mRNA            EST       17-JUL-2000
DEFINITION  sb50d01.y1 Gm-c1011 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
            Gm-c1011-314 5' similar to TR:Q26195 Q26195 PVA1 GENE. ;, mRNA
            sequence.
ACCESSION   AI748274
VERSION     AI748274.1  GI:5126538
KEYWORDS    EST.
SOURCE      soybean.
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 64)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
            ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
            Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
            ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
            ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
            ,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available through: Genome Systems, Inc. 4633 World
            Parkway Circle St. Louis, Missouri 63134 For further information
            call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314)
            427-3324 or contact: clones@genomesystems.com or
            info@genomesystems.com web site: www.genomesystems.com
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Insert Length: 482   Std Error: 0.00
            High quality!sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .64
                     /organism="Glycine max"
                     /db_xref="taxon:3847"
                     /clone="GENOME SYSTEMS CLONE ID: Gm-c1011-314"
                     /clone_lib="Gm-c1011"
                     /tissue_type="immature cotyledons of greenhouse grown
                     plants"
                     /lab_host="DH10B"
                     /note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
                     XhoI; This cDNA library was constructed from mRNA isolated
                     from immature cotyledons (100-200mgs) of greenhouse grown
                     plants. The cDNA library was prepared using the Life
                     Technologies pSuperScript cDNA library construction kit.
                     Complementary DNA was synthesized from mRNA using a poly
                     (dT) sequence with a Not I restrictions site. Sal I
                     linkers adapters were ligated to the blunt-ended cDNA
                     fragments followed by NotI digestion. The cDNA fragments
                     were directionally cloned into the NotI-Sal I restriction
                     site of the pSPORT 1 vector. The ligated cDNA fragments
                     were transformed into E. coli ElectroMax DH10B host cells.
                     This library was constructed by Dr. Lila Vodkin and Dr.
                     Anu Khanna."
BASE COUNT       8 a     25 c      6 g     25 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 64;
  Best Local Similarity   100.0%;  Pred. No. 3.6e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      46 GACGTTCG 53


RESULT  18
AA458519
LOCUS       AA458519       64 bp    mRNA            EST       09-JUN-1997
DEFINITION  zx96b04.r1 Soares ovary tumor NbHOT Homo sapiens cDNA clone
            IMAGE:811567 5' similar to gb:Z37336_cds1 IG KAPPA CHAIN V-I REGION
            (HUMAN);, mRNA sequence.
ACCESSION   AA458519
VERSION     AA458519.1  GI:2183426
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 64)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.
            , Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie
            ,T., Waterston,R. and Wilson,R.
  TITLE     WashU-Merck EST Project 1997
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -28m13 rev2 ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .64
                     /organism="Homo sapiens"
                     /db_xref="GDB:6042479"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:811567"
                     /clone_lib="Soares ovary tumor NbHOT"
                     /sex="Female"
                     /tissue_type="ovarian tumor"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: ovary; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer (5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library constructed by Bento Soares and
                     M.Fatima Bonaldo. "
BASE COUNT      16 a     16 c     20 g     12 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 64;
  Best Local Similarity   100.0%;  Pred. No. 3.6e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      22 GACGTTCG 29
```

```
RESULT  19
AA606766
LOCUS       AA606766       64 bp    mRNA            EST       30-SEP-1997
DEFINITION  vm86e04.r1 Knowles Solter mouse blastocyst B1 Mus musculus cDNA
            clone IMAGE:1005150 5' similar to SW:NUGM_MARPO P34944 PROBABLE
            NADH-UBIQUINONE OXIDOREDUCTASE SUBUNIT ;, mRNA sequence.
ACCESSION   AA606766
VERSION     AA606766.1  GI:2455659
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 64)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:569366
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .64
                     /organism="Mus musculus"
                     /strain="B6D2 F1/J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1005150"
                     /clone_lib="Knowles Solter mouse blastocyst B1"
                     /tissue_type="blastocyst"
                     /dev_stage="embryo (pre-implantation)"
                     /lab_host="DH10B"
                     /note="Organ: embryo; Vector: pSPORT; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally from mRNA prepared
                     from 800 blastocysts.  Primer: SalI(dT):
                     5'-CGGTCGACCGTCGACCGTTTTTTTTTTTTTT-3'.  cDNAs  were
                     cloned into the NotI/SalI sites of a pSPORT vector (Life
                     Technologies).  Two different size selections:  B1 (larger
                     inserts) and B3."
BASE COUNT       16 a     18 c     17 g     13 t
ORIGIN

  Query Match             100.0%;  Score 8;  DB 10;  Length 64;
  Best Local Similarity   100.0%;  Pred. No. 3.6e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db     47 GACGTTCG 54


RESULT  20
H53706
LOCUS       H53706         64 bp    mRNA            EST       20-SEP-1995
DEFINITION  yu38c10.r1 Soares ovary tumor NbHOT Homo sapiens cDNA clone
            IMAGE:236082 5' similar to gb:L11700 IG KAPPA CHAIN V-IV REGION
            (HUMAN);, mRNA sequence.
ACCESSION   H53706
VERSION     H53706.1  GI:993853
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 64)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
            ,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
            ,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
            Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E.,
            Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
  MEDLINE   97044478
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 754
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 754    Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 239.
FEATURES             Location/Qualifiers
     source          1. .64
                     /organism="Homo sapiens"
                     /db_xref="GDB:3863001"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:236082"
                     /clone_lib="Soares ovary tumor NbHOT"
                     /sex="Female"
                     /tissue_type="ovarian tumor"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: ovary; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library constructed by Bento Soares and
                     M.Fatima Bonaldo. "
BASE COUNT       20 a     14 c     17 g     13 t
ORIGIN

  Query Match             100.0%;  Score 8;  DB 11;  Length 64;
  Best Local Similarity   100.0%;  Pred. No. 3.6e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db     22 GACGTTCG 29


RESULT  21
TA303C12Q
LOCUS       TA303C12Q      65 bp    DNA             GSS       13-DEC-2000
DEFINITION  T. brucei sheared genomic DNA clone 303c12, reverse sequence,
            genomic survey sequence.
ACCESSION   AL489246
VERSION     AL489246.1  GI:11864608
KEYWORDS    GSS.
SOURCE      Trypanosoma brucei.
  ORGANISM  Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
```

```
REFERENCE   1  (bases 1 to 65)
  AUTHORS   Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
            Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
            Melville,S.E., Rajandream,M.A. and Barrell,B.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
            project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
            Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
            nhl@sanger.ac.uk
COMMENT     Constructed at the Institute for Genomic Research (TIGR),
            Rockville, MD. Genomic DNA isolated from a cloned population of
            Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
            to give a tight size distribution (
            4 kb).  The v + i method used for the library construction is
            described in detail in Smith, H. and Venter, J.C. (Making small
            insert libraries for whole genome shotgun sequencing projects.  In
            Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
            Barrell, Oxford University Press, 1999).
            Email: nelsayed@tigr.org
            Details of T. brucei sequencing at the Sanger Centre are available
            at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES             Location/Qualifiers
     source          1. .65
                     /organism="Trypanosoma brucei"
                     /strain="TREU927"
                     /db_xref="taxon:5691"
                     /clone="303c12"
BASE COUNT       16 a     10 c     22 g     17 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 13;  Length 65;
  Best Local Similarity   100.0%;  Pred. No. 3.6e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db      8 GACGTTCG 15


RESULT  22
AW063368/c
LOCUS       AW063368       66 bp    mRNA            EST       07-DEC-2000
DEFINITION  TN0743 KRIBB Human TN intrathymic T-cell cDNA library Homo sapiens
            cDNA 3', mRNA sequence.
ACCESSION   AW063368
VERSION     AW063368.1  GI:8887305
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 66)
  AUTHORS   Goh,S.-H., Park,J.-H., Lee,Y.J., Lee,H.G., Yoo,H.-S., Lee,I.-C.,
            Park,J.-H., Kim,Y.-S. and Lee,C.-C.
  TITLE     Gene expression profile and identification of differentially
            expressed transcripts during human intrathymic T-cell development
            by cDNA sequencing analysis
  JOURNAL   Genomics 70 (1), 1-18 (2000)
COMMENT     Contact: Sung-Ho Goh
            Genome Center
            Korea Research Institute of Bioscience and Biotechnology
            Oun-dong 52, Yu Sung-Gu, Daejon 305-333, Republic of Korea
            Tel: 82-42-860-4473
            Fax: 82-42-860-4479
            Email: gohsh@mail.kribb.re.kr
            Seq primer: T7
            High quality sequence stop: 66
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .66
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="KRIBB Human TN intrathymic T-cell cDNA
                     library"
                     /tissue_type="Thymus"
                     /cell_type="Intrathymic T-cell"
                     /dev_stage="CD3-4-8- triple negative stage"
                     /note="Vector: pGEM-T; cDNA was made from total
                     cytoplasmic RNA of sorted human intrathymic CD3-4-8-
                     T-cell, adaptor ligated, amplified with PCR, and cloned
                     into pGEM-T vector."
BASE COUNT       11 a     22 c     23 g     10 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 66;
  Best Local Similarity   100.0%;  Pred. No. 3.6e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db     56 GACGTTCG 49


RESULT  23
AA688966/c
LOCUS       AA688966       67 bp    mRNA            EST       12-DEC-1997
DEFINITION  vs01a10.r1 Barstead mouse irradiated colon MPLRB7 Mus musculus cDNA
            clone IMAGE:1136922 5' similar to SW:RL2B_HUMAN P29316 60S
            RIBOSOMAL PROTEIN L23A. [2] SW:RL2B_HUMAN ;, mRNA sequence.
ACCESSION   AA688966
VERSION     AA688966.1  GI:2678395
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 67)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:618194
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .67
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1136922"
                     /clone_lib="Barstead mouse irradiated colon MPLRB7"
                     /dev_stage="8 weeks"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: EcoRI; Site_2: NotI; Tissue obtained
                     from 8 week old mouse.  Colon was harvested 72 hours after
                     irradiation with 1400 Gys.  1st strand cDNA was primed
                     with  a Not I - oligo(dT) primer
                     [5'TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTT
                     T 3'];  double-stranded cDNA was ligated to Eco RI
```

```
                adaptors [AATTCGGATCCTTG], digested with Not I and cloned
                into the Not I and Eco RI sites of the modified pT7T3
                vector. Library constructed by Bob Barstead. "
BASE COUNT       21 a     20 c     19 g      7 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 67;
  Best Local Similarity   100.0%;  Pred. No. 3.6e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db     30 GACGTTCG 23


RESULT  24
AI956572/c
LOCUS       AI956572       70 bp    mRNA            EST       20-AUG-1999
DEFINITION  ul78h08.y1 Sugano mouse kidney mkia Mus musculus cDNA clone
            IMAGE:2136735 5' similar to TR:O14597 O14597 NON-FUNCTIONAL FOLATE
            BINDING PROTEIN. ;, mRNA sequence.
ACCESSION   AI956572
VERSION     AI956572.1  GI:5749281
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 70)
  AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
            ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
            ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     The WashU-NCI Mouse EST Project 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:1001411
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: custom primer used
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .70
                     /organism="Mus musculus"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:2136735"
                     /clone_lib="Sugano mouse kidney mkia"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: kidney; Vector: pME18S-FL3; Site_1: DraIII
                     (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
                     was primed with an oligo(dT) primer
                     [ATGTGGCCTTTTTTTTTTTTTTTT]; double-stranded cDNA was
                     ligated to a  DraIII adaptor [TGTTGGCCTACTGG], digested
                     and cloned into distinct DraIII sites of the pME18S-FL3
                     vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
                     be used to isolate the cDNA insert.  Size selection was
                     performed to exclude fragments <1.5kb.  Library
                     constructed by Dr. Sumio Sugano (University of Tokyo
                     Institute of Medical Science).  Custom primers  for
                     sequencing: 5' end primer CTTCTGCTCTAAAAGCTGCG and 3' end
                     primer CGACCTGCAGCTCGAGCACA."
BASE COUNT       12 a     24 c     17 g     17 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 70;
  Best Local Similarity   100.0%;  Pred. No. 3.6e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db     41 GACGTTCG 34


RESULT  25
AI986743/c
LOCUS       AI986743       70 bp    mRNA            EST       10-MAY-2001
DEFINITION  rs16h09.y1 Sommer Pristionchus Pristionchus pacificus cDNA 5', mRNA
            sequence.
ACCESSION   AI986743
VERSION     AI986743.1  GI:5815898
KEYWORDS    EST.
SOURCE      Pristionchus pacificus.
  ORGANISM  Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE   1  (bases 1 to 70)
  AUTHORS   McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T.
            , Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y.
            , Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagareishvili,R.
            , Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe
            ,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S.
            , Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and
            Wilson,R.
  TITLE     The Washington Univ. Nematode EST Project, 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: McCarter JP
            The Washington Univ. Nematode EST Project, 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            The library was constructed by Dr. Ralf Sommer DNA Sequencing by:
            Washington University Genome Sequencing Center
             Contact Dr. Ralf Sommer (ralf.sommer@tuebingen.mpg.de) for
            information about this clone.
            Seq primer: -40RP from Gibco
            High quality sequence stop: 62.
FEATURES             Location/Qualifiers
     source          1. .70
                     /organism="Pristionchus pacificus"
                     /strain="PS 312"
                     /db_xref="taxon:54126"
                     /clone_lib="Sommer Pristionchus"
                     /sex="predominantly hermaphroditic"
                     /dev_stage="mixed stages (embryo to adult)"
                     /lab_host="not applicable (host cell line)"
                     /note="Vector: Uni-ZAP XR Vector (Stratagene); Site_1: 5'
                     EcoRI; Site_2: 3' XhoI; 1st strand cDNA was primed with a
                     XhoI - oligo(dT) primer. Double-stranded cDNA was ligated
                     to EcoRI adaptors digested with XhoI and cloned into XhoI
                     and EcoRI sites. Primary complexity of the library was 10
                     in the 7th. The library went through one round of
                     amplification."
BASE COUNT       18 a     15 c     15 g     22 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 70;
  Best Local Similarity   100.0%;  Pred. No. 3.6e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
```

```
          ||||||||
Db      47 GACGTTCG 40


RESULT  26
AI365158
LOCUS       AI365158       71 bp    mRNA            EST       15-FEB-1999
DEFINITION  qx97a08.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:2010422 3',
            mRNA sequence.
ACCESSION   AI365158
VERSION     AI365158.1  GI:4124847
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 71)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
            R. Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 407   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 67.
FEATURES             Location/Qualifiers
     source          1. .71
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2010422"
                     /clone_lib="NCI_CGAP_GC6"
                     /tissue_type="pooled germ cell tumors"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA
                     from the normalized library NCI_CGAP_GC4 was prepared, and
                     ss circles were made in vitro. Following HAP purification,
                     this DNA was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from a pool
                     of 5,000 clones made from the same library (cloneIDs
                     1257096-1258631, 1469064-1470983, and 1475592-1476743).
                     Subtraction by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT       15 a     20 c     14 g     22 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 71;
  Best Local Similarity   100.0%;  Pred. No. 3.6e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db      7 GACGTTCG 14


RESULT  27
AA243275/c
LOCUS       AA243275       72 bp    mRNA            EST       11-MAR-1998
DEFINITION  zr26g11.r1 Stratagene NT2 neuronal precursor 937230 Homo sapiens
            cDNA clone IMAGE:664580 5', mRNA sequence.
ACCESSION   AA243275
VERSION     AA243275.1  GI:1874085
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 72)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 2343   Std Error: 0.00
            Seq primer: -28m13 rev1 ET from Amersham
            High quality sequence stop: 57.
FEATURES             Location/Qualifiers
     source          1. .72
                     /organism="Homo sapiens"
                     /db_xref="GDB:5426609"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:664580"
                     /clone_lib="Stratagene NT2 neuronal precursor 937230"
                     /tissue_type="neuroepithelial cells"
                     /dev_stage="Ntera-2 neuroepithelial cells"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Organ: brain; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                     Oligo dT. Uninduced, exponentially growing neuroepithelial
                     cells (Ntera-2/c1.D1). Average insert size: 1.0 kb;
                     Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCACGAG
                     3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT       12 a     21 c     27 g     12 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 72;
  Best Local Similarity   100.0%;  Pred. No. 3.6e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db      42 GACGTTCG 35


RESULT  28
AA967742/c
LOCUS       AA967742       73 bp    mRNA            EST       19-MAY-1998
DEFINITION  uh04c05.r1 Soares mouse hypothalamus NMHy Mus musculus cDNA clone
            IMAGE:1616936 5' similar to TR:Q29269 Q29269 UNKNOWN PROTEIN ;,
            mRNA sequence.
ACCESSION   AA967742
VERSION     AA967742.1  GI:3141635
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 73)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
```

```
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:956236
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .73
                      /organism="Mus musculus"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:1616936"
                      /clone_lib="Soares mouse hypothalamus NMHy"
                      /tissue_type="hypothalamus"
                      /lab_host="DH10B"
                      /note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
                      modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                      strand cDNA was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCCAAGGTTTTTTTTTTTTTTTTTTTTTTT
                      T 3']; double-stranded cDNA was ligated to Eco RI adaptors
                      (Pharmacia), digested with Not I and cloned into the Not I
                      and Eco RI sites of the modified pT7T3 vector. RNA
                      provided by Dr. Wolfgang Liedtke.  Library went through
                      two rounds of normalization, and was constructed by Bento
                      Soares and M.Fatima Bonaldo."
BASE COUNT       19 a     24 c     18 g     12 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 73;
  Best Local Similarity   100.0%;  Pred. No. 3.6e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      73 GACGTTCG 66


RESULT  29
AI180756/c
LOCUS       AI180756       73 bp    mRNA            EST       08-OCT-1998
DEFINITION  ub91f11.r1 Soares_mammary_gland_NbMMG Mus musculus cDNA clone
            IMAGE:1395885 5' similar to TR:Q29269 Q29269 UNKNOWN PROTEIN ;,
            mRNA sequence.
ACCESSION   AI180756
VERSION     AI180756.1  GI:3731394
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 73)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:907601
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .73
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:1395885"
                      /clone_lib="Soares_mammary_gland_NbMMG"
                      /sex="male"
                      /tissue_type="mammary gland"
                      /dev_stage="4 weeks"
                      /lab_host="DH10B"
                      /note="Organ: mammary gland; Vector: pT7T3D-Pac (Pharmacia
                      ) with a modified polylinker; Site_1: Not I; Site_2: Eco
                      RI; 1st strand cDNA was primed with a Not I - oligo(dT)
                      primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCCAATGGTTTTTTTTTTTTTTTTTTTTTT
                      T 3']; double-stranded cDNA was ligated to Eco RI
                      adaptors (Pharmacia), digested with Not I and cloned into
                      the Not I and Eco RI sites of the modified pT7T3 vector.
                      RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
                      constructed and normalized by Bento Soares and M.Fatima
                      Bonaldo."
BASE COUNT       19 a     25 c     18 g     11 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 73;
  Best Local Similarity   100.0%;  Pred. No. 3.6e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      73 GACGTTCG 66


RESULT  30
AA120541/c
LOCUS       AA120541       73 bp    mRNA            EST       19-NOV-1996
DEFINITION  mn12c03.r1 Beddington mouse embryonic region Mus musculus cDNA
            clone IMAGE:537700 5' similar to TR:G348688 G348688
            BETA-GALACTOSIDASE ALPHA-PEPTIDE ;, mRNA sequence.
ACCESSION   AA120541
VERSION     AA120541.1  GI:1675669
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 73)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:324636
            Trace considered overall poor quality
```

```
            Possible reversed clone: similarity on wrong strand
            Seq primer: -40m13 ET
            High quality sequence stop: 1.
FEATURES                Location/Qualifiers
     source             1. .73
                        /organism="Mus musculus"
                        /strain="C57BL6 x DBA"
                        /db_xref="taxon:10090"
                        /clone="IMAGE:537700"
                        /clone_lib="Beddington mouse embryonic region"
                        /sex="pooled"
                        /tissue_type="embryo"
                        /dev_stage="7.5dpc"
                        /lab_host="DH12S"
                        /note="Organ: whole embryo; Vector: pCMV-SPORT; Site_1:
                        SalI; Site_2: NotI; Cloned unidirectionally.  Primer:
                        Oligo dT. Gastrulating embryos were collected at 7.5dpc
                        from C57BL6 x DBA matings, excluding embryos that had
                        developed head folds and all extraembryonic tissues.
                        Average insert size: 1.3 kb (range: 0.5 - 3.0 kb).
                        Referenced in Development 121, 2479-2489 (1995)"
BASE COUNT       14 a      22 c      25 g      12 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 73;
  Best Local Similarity   100.0%;  Pred. No. 3.6e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      11 GACGTTCG 4


RESULT  31
AA863599/c
LOCUS       AA863599       76 bp    mRNA            EST       11-MAR-1998
DEFINITION  vx05e04.r1 Soares_thymus_2NbMT Mus musculus cDNA clone
            IMAGE:1263582 5' similar to TR:Q29269 Q29269 UNKNOWN PROTEIN ;,
            mRNA sequence.
ACCESSION   AA863599
VERSION     AA863599.1  GI:2956078
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 76)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:666134
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 1.
FEATURES                Location/Qualifiers
     source             1. .76
                        /organism="Mus musculus"
                        /strain="C57BL/6J"
                        /db_xref="taxon:10090"
                        /clone="IMAGE:1263582"
                        /clone_lib="Soares_thymus_2NbMT"
                        /sex="male"
                        /tissue_type="Thymus"
                        /dev_stage="4 weeks"
                        /lab_host="DH10B"
                        /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                        polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                        was primed with a Not I - oligo(dT) primer [5'
                        TGTTACCAATCTGAAGTGGGAGCGGCCGCGTTTTTTTTTTTTTTTTTTTTTTTT
                        3']; double-stranded cDNA was ligated to Eco RI adaptors
                        (Pharmacia), digested with Not I and cloned into the Not I
                        and Eco RI sites of the modified pT7T3 vector. RNA
                        provided by Dr. Bertrand Jordan. Library went through two
                        rounds of normalization, and was constructed by Bento
                        Soares and M.Fatima Bonaldo."
BASE COUNT       19 a      25 c      19 g      13 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 76;
  Best Local Similarity   100.0%;  Pred. No. 3.7e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      76 GACGTTCG 69


RESULT  32
AW717722/c
LOCUS       AW717722       78 bp    mRNA            EST       19-APR-2000
DEFINITION  i1b08nm.f2 Neurospora crassa morning cDNA library Neurospora crassa
            cDNA clone i1b08nm 5', mRNA sequence.
ACCESSION   AW717722
VERSION     AW717722.1  GI:7606973
KEYWORDS    EST.
SOURCE      Neurospora crassa.
  ORGANISM  Neurospora crassa
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
            Sordariales; Sordariaceae; Neurospora.
REFERENCE   1  (bases 1 to 78)
  AUTHORS   Zhu,H., Lai,H., Kupfer,D., Dunlap,J.C. and Roe,B.A.
  TITLE     Two Neurospora crassa EST Databases
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
            Department of Chemistry and Biochemistry
            Advanced Center for Genome Technology, University of Oklahoma
            620 Parrington Oval, Norman, OK 73019, USA
            Tel: 405 325 4912
            Fax: 405 325 7762
            Email: broe@ou.edu
            We anticipate the future release of the cDNA clones to the Fungal
            Genetics Stock Center
            Seq primer: Universal Forward Primer
            High quality sequence stop: 68.
FEATURES                Location/Qualifiers
     source             1. .78
                        /organism="Neurospora crassa"
                        /strain="bd, frq7 A"
                        /db_xref="taxon:5141"
                        /clone="i1b08nm"
                        /clone_lib="Neurospora crassa morning cDNA library"
                        /tissue_type="tissue harvested following 22hr growth in
                        dark"
                        /note="Vector: pBlueScript SK-; Site_1: XbaI; Site_2:
                        EcoRI; See: Bell-Pedersen,D., et al. PNAS 93:13096,1996.
                        5' end of cDNA cloned into XbaI site of pBluescript; 3'
                        end of cDNA cloned into EcoRI site of pBluescript"
BASE COUNT       19 a      26 c      16 g      17 t
ORIGIN
```

```
  Query Match             100.0%;  Score 8;  DB 10;  Length 78;
  Best Local Similarity   100.0%;  Pred. No. 3.7e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      28 GACGTTCG 21


RESULT  33
AA284594
LOCUS       AA284594       86 bp    mRNA            EST       08-AUG-1997
DEFINITION  zt22b11.r1 Soares ovary tumor NbHOT Homo sapiens cDNA clone
            IMAGE:713853 5' similar to gb:Z37336_cds1 IG KAPPA CHAIN V-I REGION
            (HUMAN);, mRNA sequence.
ACCESSION   AA284594
VERSION     AA284594.1  GI:1927505
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 86)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
            ,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
            ,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
            Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E.,
            Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
  MEDLINE   97044478
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1009   Std Error: 0.00
            Seq primer: -28m13 rev2 ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .86
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:713853"
                     /clone_lib="Soares ovary tumor NbHOT"
                     /sex="Female"
                     /tissue_type="ovarian tumor"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: ovary; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library constructed by Bento Soares and
                     M.Fatima Bonaldo. "
BASE COUNT       26 a     21 c     21 g     18 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 86;
  Best Local Similarity   100.0%;  Pred. No. 3.8e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      53 GACGTTCG 60


RESULT  34
AA239711
LOCUS       AA239711       91 bp    mRNA            EST       03-MAR-1997
DEFINITION  my15e08.r1 Barstead mouse heart MPLRB3 Mus musculus cDNA clone
            IMAGE:695942 5', mRNA sequence.
ACCESSION   AA239711
VERSION     AA239711.1  GI:1863733
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 91)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:429502
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 77.
FEATURES             Location/Qualifiers
     source          1. .91
                     /organism="Mus musculus"
                     /strain="BALB/c"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:695942"
                     /clone_lib="Barstead mouse heart MPLRB3"
                     /sex="mixed"
                     /tissue_type="heart"
                     /dev_stage="6 weeks"
                     /lab_host="DH10B"
                     /note="Organ: heart; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     [CTTGGATTCGGTACC], digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     Library constructed by Bob Barstead."
BASE COUNT       16 a     24 c     30 g     21 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 91;
  Best Local Similarity   100.0%;  Pred. No. 3.8e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      20 GACGTTCG 27


RESULT  35
AA634931
LOCUS       AA634931       92 bp    mRNA            EST       06-MAR-1998
DEFINITION  ab28h02.r1 Stratagene lung (#937210) Homo sapiens cDNA clone
            IMAGE:842163 5' similar to gb:L03555 IG KAPPA CHAIN PRECURSOR V-I
            REGION (HUMAN);, mRNA sequence.
ACCESSION   AA634931
VERSION     AA634931.1  GI:2558145
```

```
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 92)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1023   Std Error: 0.00
            Seq primer: -28m13 rev1 ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .92
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:842163"
                     /clone_lib="Stratagene lung (#937210)"
                     /sex="male"
                     /dev_stage="72 years"
                     /lab_host="SOLR cells (kanamycin resistant)"
                     /note="Organ: lung; Vector: pBluescript SK-; Site_1: EcoRI
                     ; Site_2: XhoI; Cloned unidirectionally.  Primer: Oligo
                     dT. normal lung.  Average insert size: 1.0 kb; Uni-ZAP XR
                     Vector; ~5' adaptor sequence: 5' GAATTCGGCACGAG 3' ~3'
                     adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTTT 3'"
BASE COUNT       25 a      18 c      24 g      24 t       1 others
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 92;
  Best Local Similarity   100.0%;  Pred. No. 3.8e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      50 GACGTTCG 57


RESULT  36
AA661504/c
LOCUS       AA661504       94 bp     mRNA             EST       03-DEC-1997
DEFINITION  nr18c12.s1 NCI_CGAP_Ew1 Homo sapiens cDNA clone IMAGE:1168342
            similar to TR:G189397 G189397 HYPOTHETICAL 33.4 KD PROTEIN. ;, mRNA
            sequence.
ACCESSION   AA661504
VERSION     AA661504.1  GI:2615595
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 94)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lee Helman, M.D., Michael R. Emmert-Buck, M.D.,
            Ph.D.
            cDNA Library Preparation: David B. Krizman, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Insert Length: 476   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .94
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1168342"
                     /clone_lib="NCI_CGAP_Ew1"
                     /tissue_type="Ewing's sarcoma"
                     /lab_host="DH10B"
                     /note="Vector: pAMP10; mRNA made from Ewing's sarcoma,
                     cDNA made by oligo-dT priming.  Non-directionally cloned.
                     Size-selected on agarose gel, average insert  size 600 bp.
                     Reference:  Krizman et al. (1996) Cancer Research
                     56:5380-5383."
BASE COUNT       30 a      34 c      19 g      11 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 94;
  Best Local Similarity   100.0%;  Pred. No. 3.8e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      59 GACGTTCG 52


RESULT  37
AI957911/c
LOCUS       AI957911       94 bp     mRNA             EST       07-JUN-2001
DEFINITION  fd08a06.x1 Zebrafish WashU MPIMG EST Danio rerio cDNA clone
            IMAGE:3730258 3' similar to SW:GATM_PIG P50441 GLYCINE
            AMIDINOTRANSFERASE ;contains element MER22 repetitive element ;,
            mRNA sequence.
ACCESSION   AI957911
VERSION     AI957911.1  GI:5750620
KEYWORDS    EST.
SOURCE      zebrafish.
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Rasborinae; Danio.
REFERENCE   1  (bases 1 to 94)
  AUTHORS   Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
            ,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
            ,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
            Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
            Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
            and Wilson,R.
  TITLE     WashU Zebrafish EST Project 1998
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Stephen L. Johnson
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: zbrafish@watson.wustl.edu
            cDNA Library Preparation: Matthew Clark. cDNA Library Arrayed by:
            Matthew Clark. DNA Sequencing by: Washington University Genome
            Sequencing Center Clone distribution: Genome Systems, St. Louis,
            Missouri (web address: www.genomesystems.com) (email contact:
            info@genomesystems.com) and Research Genetics, Huntsville, Alabama
            (web address: www.resgen.com) (email contact: info@resgen.com) and
            RessourcenZentrumPrimarDatenbank, Berlin, Germany (web address:
            www.rzpd.de)
```

```
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: T7 ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .94
                     /organism="Danio rerio"
                     /db_xref="taxon:7955"
                     /clone="IMAGE:3730258"
                     /clone_lib="Zebrafish WashU MPIMG EST"
                     /sex="mixed"
                     /tissue_type="26 somite embryos, adult livers, shield
                     stage embryos"
                     /lab_host="XL1-blue MRF"
                     /note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI; 1st
                     strand cDNA was primed with a Not I - oligo(dT)15 primer
                     [5'pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT3']];
                     double-stranded cDNA was ligated to Sal I adaptors (BRL),
                     digested with Not I and cloned into the Not I and Sal I
                     sites of the pSPORT1 vector (BRL). Library was constructed
                     by Matthew Clark (Lehrach lab; ICRF, London and Max Planck
                     Institut fuer Molekulare Genetik,Berlin). cDNAs for EST
                     analysis were selected following oligonucleotide
                     hybridization fingerprinting of arrayed clones from
                     zebrafish late somitogenesis (26 ss), adult liver or
                     embryonic shield stage (5.6 h) libraries.  Fingerprint
                     data were used to computationally cluster cDNAs, and a
                     single cDNA from each cluster was chosen for sequencing.
                     In some cases multiple members of the same cluster were
                     sequenced to assess clustering parameters or single clones
                     were sequenced additional times to assess quality
                     control."
BASE COUNT       22 a     28 c     27 g     17 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 94;
  Best Local Similarity   100.0%;  Pred. No. 3.8e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      34 GACGTTCG 27


RESULT  38
AA722920/c
LOCUS       AA722920       95 bp    mRNA            EST       02-JAN-1998
DEFINITION  zg81c01.s1 Soares_fetal_heart_NbHH19W Homo sapiens cDNA clone
            IMAGE:399744 3' similar to gb:D10522 Human mRNA for 80K-L protein,
            complete cds. (HUMAN);, mRNA sequence.
ACCESSION   AA722920
VERSION     AA722920.1  GI:2740627
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 95)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Possible reversed clone: polyT not found
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 90.
FEATURES             Location/Qualifiers
     source          1. .95
                     /organism="Homo sapiens"
                     /db_xref="GDB:1307567"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:399744"
                     /clone_lib="Soares_fetal_heart_NbHH19W"
                     /sex="unknown"
                     /dev_stage="19 weeks"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATCTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by
                     M.Fatima Bonaldo. This library was constructed from the
                     same fetus as the fetal lung library, Soares fetal lung
                     NbHL19W."
BASE COUNT       17 a     29 c     37 g     12 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 95;
  Best Local Similarity   100.0%;  Pred. No. 3.8e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACGTTCG 8
           ||||||||
Db      19 GACGTTCG 12


RESULT  39
AW630585
LOCUS       AW630585       95 bp    mRNA            EST       31-MAR-2000
DEFINITION  hh81g04.y1 NCI_CGAP_GU1 Homo sapiens cDNA clone IMAGE:2969238 5'
            similar to gb:Z37336_cds1 IG KAPPA CHAIN V-I REGION (HUMAN);, mRNA
            sequence.
ACCESSION   AW630585
VERSION     AW630585.1  GI:7377375
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 95)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Other_ESTs: hh81g04.x1
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Chris Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Life
            Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The
            I.M.A.G.E. Consortium DNA Sequencing by: Washington University
            Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            image.llnl.gov/image/html/iresources.shtml
            Seq primer: -40RP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .95
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2969238"
```

```
                     /clone_lib="NCI_CGAP_GU1"
                     /tissue_type="2 pooled high-grade transitional cell
                     tumors"
                     /lab_host="DH10B"
                     /note="Organ: genitourinary tract; Vector: pCMV-SPORT6;
                     Site_1: SalI; Site_2: NotI; Cloned unidirectionally.
                     Primer: Oligo dT.   Library constructed by Life
                     Technologies."
BASE COUNT       28 a     25 c     21 g     21 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 95;
  Best Local Similarity   100.0%;  Pred. No. 3.8e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db     53 GACGTTCG 60


RESULT  40
AA426003
LOCUS       AA426003       95 bp    mRNA            EST       16-OCT-1997
DEFINITION  zw17e07.r1 Soares ovary tumor NbHOT Homo sapiens cDNA clone
            IMAGE:769572 5' similar to gb:Z37336_cds1 IG KAPPA CHAIN V-I REGION
            (HUMAN);, mRNA sequence.
ACCESSION   AA426003
VERSION     AA426003.1  GI:2107879
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 95)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.
            , Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie
            ,T., Waterston,R. and Wilson,R.
  TITLE     WashU-Merck EST Project 1997
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -28m13 rev2 ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .95
                     /organism="Homo sapiens"
                     /db_xref="GDB:5979442"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:769572"
                     /clone_lib="Soares ovary tumor NbHOT"
                     /sex="Female"
                     /tissue_type="ovarian tumor"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: ovary; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library constructed by Bento Soares and
                     M.Fatima Bonaldo. "
BASE COUNT       30 a     23 c     21 g     21 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 95;
  Best Local Similarity   100.0%;  Pred. No. 3.8e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db     53 GACGTTCG 60


RESULT  41
AI329158
LOCUS       AI329158       98 bp    mRNA            EST       28-DEC-1998
DEFINITION  b1g10ne.r1 Neurospora crassa evening cDNA library Neurospora crassa
            cDNA clone b1g10ne 3', mRNA sequence.
ACCESSION   AI329158
VERSION     AI329158.1  GI:4065717
KEYWORDS    EST.
SOURCE      Neurospora crassa.
  ORGANISM  Neurospora crassa
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
            Sordariales; Sordariaceae; Neurospora.
REFERENCE   1  (bases 1 to 98)
  AUTHORS   Zhu,H., Lai,H., Kupfer,D., Dunlap,J.C. and Roe,B.A.
  TITLE     Two Neurospora crassa EST Databases
  JOURNAL   Unpublished (1998)
COMMENT     Other_ESTs: b1g10ne.f1
            Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
            Department of Chemistry and Biochemistry
            Advanced Center for Genome Technology, University of Oklahoma
            620 Parrington Oval, Norman, OK 73019, USA
            Tel: 405 325 4912
            Fax: 405 325 7762
            Email: broe@ou.edu
            We anticipate the future release of the cDNA clones to the Fungal
            Genetics Stock Center
            Possible reversed clone: polyT not found
            Seq primer: Universal Reverse Primer
            High quality sequence stop: 59.
FEATURES             Location/Qualifiers
     source          1. .98
                     /organism="Neurospora crassa"
                     /strain="Strain 30-7 (bd; A)"
                     /db_xref="taxon:5141"
                     /clone="b1g10ne"
                     /clone_lib="Neurospora crassa evening cDNA library"
                     /tissue_type="tissue harvested following 22hr growth in
                     dark"
                     /note="Vector: pBlueScript SK-; Site_1: XbaI; Site_2:
                     EcoRI; See: Bell-Perdersen,D., et al. PNAS 93:13096,1996.
                     5' end of cDNA cloned into XbaI site of pBluescript; 3'
                     end of cDNA cloned into EcoRI site of pBluescript"
BASE COUNT       25 a     20 c     32 g     21 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 98;
  Best Local Similarity   100.0%;  Pred. No. 3.9e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db     27 GACGTTCG 34


RESULT  42
AA593996
LOCUS       AA593996       99 bp    mRNA            EST       25-SEP-1997
DEFINITION  nn16f03.s1 NCI_CGAP_Co12 Homo sapiens cDNA clone IMAGE:1084061 3'
            similar to gb:L11700 IG KAPPA CHAIN V-IV REGION (HUMAN);, mRNA
            sequence.
ACCESSION   AA593996
VERSION     AA593996.1  GI:2409346
KEYWORDS    EST.
```

```
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Stratagene, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 2427   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1084061"
                     /clone_lib="NCI_CGAP_Co12"
                     /sex="mixed"
                     /tissue_type="colon tumor"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Organ: colon; Vector: Bluescript SK-; Site_1: EcoRI
                     ; Site_2: XhoI; Cloned unidirectionally.  Primer: Oligo
                     dT. Pooled colon tumors. 5' adaptor sequence: 5'
                     GAATTCGGCACGAG 3' 3' adaptor sequence: 5'
                     CTCGAGTTTTTTTTTTTTTTTTT 3' Average insert size:  1.2 kb."
BASE COUNT       28 a     24 c     22 g     25 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 99;
  Best Local Similarity   100.0%;  Pred. No. 3.9e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db     29 GACGTTCG 36


RESULT  43
AI622446/c
LOCUS       AI622446      100 bp    mRNA            EST       22-APR-1999
DEFINITION  486055C03.x3 486 - leaf primordia cDNA library from Hake lab Zea
            mays cDNA, mRNA sequence.
ACCESSION   AI622446
VERSION     AI622446.1  GI:4647371
KEYWORDS    EST.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 100)
  AUTHORS   Walbot,V.
  TITLE     Maize ESTs from various cDNA libraries sequenced at Stanford
            University
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 486055  row: C  column: 03.
FEATURES             Location/Qualifiers
     source          1. .100
                     /organism="Zea mays"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone_lib="486 - leaf primordia cDNA library from Hake
                     lab"
                     /tissue_type="leaf primordia"
                     /dev_stage="P7-P11 leaf"
                     /lab_host="E.coli XL1-Blue MFR'"
                     /note="Organ: shoot; Vector: Lambda zap; Hake lab cDNA
                     library."
BASE COUNT       24 a     38 c     17 g     21 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 10;  Length 100;
  Best Local Similarity   100.0%;  Pred. No. 3.9e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db     65 GACGTTCG 58


RESULT  44
BG272807
LOCUS       BG272807      100 bp    mRNA            EST       20-FEB-2001
DEFINITION  nah90g06.x1 NCI_CGAP_HN19 Homo sapiens cDNA clone IMAGE:4257994
            similar to SW:KV3M_HUMAN P18136 IG KAPPA CHAIN V-III REGION HIC
            PRECURSOR. ;, mRNA sequence.
ACCESSION   BG272807
VERSION     BG272807.1  GI:12982333
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 100)
  AUTHORS   NCI/NIDR-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute / National Institute of Dental Research,
            Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            cDNA Library Preparation: D. Krizman, Ph.D.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .100
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4257994"
                     /clone_lib="NCI_CGAP_HN19"
                     /tissue_type="normal epithelium"
                     /lab_host="DH10B"
                     /note="Organ: nasopharynx; Vector: pAMP10; mRNA made from
                     normal nasopharyngeal epithelium, cDNA made by oligo-dT
                     priming. Non-directionally cloned into UDG sites.
                     Size-selected on  agarose gel, average insert size 500 bp.
                     Primary library.  cDNA Library Preparation: David B.
                     Krizman, Ph.D. REFERENCE:  Krizman et al. (1996) Cancer
                     Research 56:5380-5383."
BASE COUNT       24 a     21 c     31 g     24 t
ORIGIN
```

```
  Query Match             100.0%;  Score 8;  DB 11;  Length 100;
  Best Local Similarity   100.0%;  Pred. No. 3.9e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db     58 GACGTTCG 65


RESULT  45
TA101B10P/c
LOCUS       TA101B10P     100 bp    DNA             GSS       13-DEC-2000
DEFINITION  T. brucei sheared genomic DNA clone 101b10, forward sequence,
            genomic survey sequence.
ACCESSION   AL458854
VERSION     AL458854.1  GI:11830896
KEYWORDS    GSS.
SOURCE      Trypanosoma brucei.
  ORGANISM  Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 100)
  AUTHORS   Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
            Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
            Melville,S.E., Rajandream,M.A. and Barrell,B.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
            project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
            Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
            nh1@sanger.ac.uk
COMMENT     Constructed at the Institute for Genomic Research (TIGR),
            Rockville, MD. Genomic DNA isolated from a cloned population of
            Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
            to give a tight size distribution (
            4 kb).  The v + i method used for the library construction is
            described in detail in Smith, H. and Venter, J.C. (Making small
            insert libraries for whole genome shotgun sequencing projects.  In
            Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
            Barrell, Oxford University Press, 1999).
            Email: nelsayed@tigr.org
            Details of T. brucei sequencing at the Sanger Centre are available
            at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES             Location/Qualifiers
     source          1. .100
                     /organism="Trypanosoma brucei"
                     /strain="TREU927"
                     /db_xref="taxon:5691"
                     /clone="101b10"
BASE COUNT       25 a     22 c     33 g     20 t
ORIGIN


  Query Match             100.0%;  Score 8;  DB 13;  Length 100;
  Best Local Similarity   100.0%;  Pred. No. 3.9e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GACGTTCG 8
          ||||||||
Db     35 GACGTTCG 28


Search completed: November 29, 2001, 14:23:50
Job time: 8083 sec
```

```
> O <
O| |O  IntelliGenetics
> O <

Quest - Quick User-directed Expression Search Tool
Release 5.4

                  -- Outline of search "papiss" --

Selected search type is key against sequence data banks or files.
Selected scope is Sequence.
Selected sequence key from "pappu445.key":
  papiss (NA) ID   papiss NA preliminary pattern
1       followed by
2         a or g
2         a or g
2         cg
2         c or t
2         c or t
2         cg

Selected files:

        File : hpvcomplete.seq

                      -- Output Parameters --

Format Options:                        File Options:
  Nucleic acid code matching   Exact    Indirect file                No
  Find non-matching hits only  No       Sequence or key file         No
  Report key used              Yes      List of hits                 Yes
  Note position of hit         Yes      Hit display                  Yes
  Display full annotations     Yes      Name and annotations         Yes
  Sequence context             50


                     -- Run Parameters --

  Run mode                     Batch
    Time to start comparison   now
    Notify at end of run       No

-------
1 match found in sequence:
 pph11 ; TOIG of: pph11  check: 3689  from: 1  to: 7931
  (from "hpvcomplete.seq")
  TOIG of: pph11  check: 3689  from: 1  to: 7931

 LOCUS       PPH11        7931 bp    DNA   circular  VRL       02-JUN-1994
 DEFINITION  Human papillomavirus type 11 (HPV-11) complete genome.
 ACCESSION   M14119
 VERSION     M14119.1  GI:333026
 KEYWORDS    complete genome.
 SOURCE      Human laryngeal papillomavirus type 11 DNA.
   ORGANISM  Human papillomavirus type 11
             Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
             Papillomavirus.
 REFERENCE   1  (bases 1 to 7931)
   AUTHORS   Dartmann,K., Schwarz,E., Gissmann,L. and Zur Hausen,H.
   TITLE     The nucleotide sequence and genome organization of human papilloma
             virus type 11
   JOURNAL   Virology 151, 124-130 (1986)
   MEDLINE   86181601
 COMMENT     ORF L1 is assumed to encode the major structural protein.
 FEATURES             Location/Qualifiers
      source          1. .7931
                      /organism="Human papillomavirus type 11"
                      /db_xref="taxon:10580"
      CAAT_signal     9. .15
                      /note="putative"
      protein_bind    35. .46
                      /note="putative"
                      /function="gene transcription"
                      /bound_moiety="E2"
      protein_bind    50. .61
                      /note="putative"
                      /function="gene transcription"
                      /bound_moiety="E2"
      TATA_signal     66. .71
                      /note="putative"
      gene            102. .554
                      /gene="E6"
      CDS             102. .554
                      /gene="E6"
                      /note="102 is position of first start codon in ORF E6;
                      putative"
                      /codon_start=1
                      /product="transforming protein"
                      /protein_id="AAA46927.1"
                      /db_xref="GI:496193"
                      /translation="MESKDASTSATSIDQLCKTFNLSLHTLQIQCVFCRNALTTAEIY
                      AYAYKNLKVVWRDNFPFAACACCLELQGKINQYRHFNYAAYAPTVEEETNEDILKVLI
                      RCYLCHKPLCEIEKLKHILGKARFIKLNNQWKGRCLHCWTTCMEDLLP"
      gene            530. .826
                      /gene="E7"
      CDS             530. .826
                      /gene="E7"
                      /note="530 is position of first start codon in ORF E7;
                      putative"
                      /codon_start=1
                      /product="transforming protein"
                      /protein_id="AAA46928.1"
                      /db_xref="GI:496194"
                      /translation="MHGRLVTLKDIVLDLQPPDPVGLHCYEQLEDSSEDEVDKVDKQD
                      AQPLTQHYQILTCCCGCDSNVRLVVECTDGDIRQLQDLLLGTLNIVCPICAPKP"
      gene            832. .2781
                      /gene="E1"
      CDS             832. .2781
                      /gene="E1"
                      /note="832 is position of first start codon in ORF E1;
                      putative"
                      /codon_start=1
                      /product="replication protein"
                      /protein_id="AAA46929.1"
                      /db_xref="GI:496195"
                      /translation="MADDSGTENEGSGCTGWFMVEAIVEHTTGTQISEDEEEEVEDSG
                      YDMVDFIDDRHITQNSVEAQALFNRQEADAHYATVQDLKRKYLGSPYVSPISNVANAV
                      ESEISPRLDAIKLTTQPKKVKRRLFETRELTDSGYGYSEVEAATQVEKHGDPENGGDG
                      QERDTGRDIEGEGVEHREAEAVDDSTREHADTSGILELLKCKDIRSTLHGKFKDCFGL
                      SFVDLIRPFKSDRTTCADWVVAGFGIHHSIADAFQKLIEPLSLYAHIQWLTNAWGMVL
                      LVLIRFKVNKSRCTVARTLGTLLNIPENHMLIEPPKIQSGVRALYWFRTGISNASTVI
                      GEAPEWITRQTVIEHSLADSQFKLTEMVQWAYDNDICEESEIAFEYAQRGDFDSNARA
                      FLNSNMQAKYVKDCAIMCRHYKHAEMKKMSIKQWIKYRGTKVDSVGNWKPIVQFLRHQ
                      NIEFIPFLSKLKLWLHGTPKKNCIAIVGPPDTGKSCFCMSLIKFLGGTVISYVNSCSH
                      FWLQPLTDAKVALLDDATQPCWTYMDTYMRNLLDGNPMSIDRKHRALTLIKCPPLLVT
                      SNIDISKEEKYKYLHSRVTTFTFPNPFPFDRNGNAVYELSDANWKCFFERLSSSLDIE
                      DSEDEEDGSNSQAFRCVPGSVVRTL"
      gene            2723. .3826
                      /gene="E2"
      CDS             2723. .3826
                      /gene="E2"
                      /note="2723 is position of first start codon in ORF E2;
                      putative"
                      /codon_start=1
                      /product="regulatory protein"
                      /protein_id="AAA46930.1"
                      /db_xref="GI:496196"
                      /translation="MEAIAKRLDACQDQLLELYEENSIDIHKHIMHWKCIRLESVLLH
                      KAKQMGLSHIGLQVVPPLTVSETKGHNAIEMQMHLESLAKTQYGVEPWTLQDTSYEMW
                      LTPPKRCFKKQGNTVEVKFDGCEDNVMEYVVWTHIYLQDNDSWVKVTSSVDAKGIYYT
                      CGQFKTYYVNFNKEAQKYGSTNHWEVCYGSTVICSPASVSSTVREVSIAEPTTYTPAQ
                      TTAPTVSACTTEDGVSAPPRKRARGPSTNNTLCVANIRSVDSTINNIVTDNYNKHQRR
                      NNCHSAATPIVQLQGDSNCLKCFRYRLNDKYKHLFELASSTWHWASPEAPHKNAIVTL
                      TYSSEEQRQQFLNSVKIPPTIRHKVGFMSLHLL"
      gene            3255. .3581
                      /gene="E4"
```

```
CDS             3255. .3581
                /gene="E4"
                /note="3255 is position of first start codon in ORF E4;
                putative"
                /codon_start=1
                /protein_id="AAA46931.1"
                /db_xref="GI:496197"
                /translation="MVVPIIGKYVMAAQLYVLLHLYLALYEKYPLLNLLHTPPHRPPP
                LQCPPAPRKTACRRRLGSEHVDRPLTTPCVWPTSDPWTVQSTTSSLTITTSTKEGTTV
                TVQLRL"
gene            3871. .4146
                /gene="E5A"
CDS             3871. .4146
                /gene="E5A"
                /note="3871 is position of first start codon in ORFE5A;
                putative"
                /codon_start=1
                /protein_id="AAA46932.1"
                /db_xref="GI:496198"
                /translation="MEVVPVQIAAATTTTLILPVVIAFAVCILSIVLIILISDFVVYT
                SVLVLTLLLYLLLWLLLTTPLQFFLLTLCVCYFPAFYIHIYIVQTQQ"
gene            4146. .4370
                /gene="E5B"
CDS             4146. .4370
                /gene="E5B"
                /note="4146 is position of first start codon in ORF E5B;
                putative"
                /codon_start=1
                /protein_id="AAA46933.1"
                /db_xref="GI:496199"
                /translation="MVMLTCHLNDGDTWLFLWLFTAFVVAVLGLLLLHYRAVHGTEKT
                KCAKCKSNRNTTVDYVYMSHGDNGDYVYMN"
polyA_signal    4371. .4376
                /note="putative"
gene            4417. .5784
                /gene="L2"
CDS             4417. .5784
                /gene="L2"
                /note="4417 is position of first start codon in ORF L2;
                putative"
                /codon_start=1
                /product="minor capsid protein"
                /protein_id="AAA46934.1"
                /db_xref="GI:496200"
                /translation="MKPRARRRKRASATQLYQTCKATGTCPPDVIPKVEHTTIADQIL
                KWGSLGVFFGGLGIGTGAGSGGRAGYIPLGSSPKPAITGGPAARPPVLVEPVAPSDPS
                IVSLIEESAIINAGAPEVVPPTQGGFTITSSESTTPAILDVSVTNHTTTSVFQNPLFT
                EPSVIQPQPPVEASGHILISAPTITSQHVEDIPLDTFVVSSSDSGPTSSTPLPRAFPR
                PRVGLYSRALQQVQVTDPAFLSTPQRLVTYDNPVYEGEDVSLQFTHESIHNAPDEAFM
                DIIRLHRPAITSRRGLVRFSRIGQRGSMYTRSGQHIGARIHYFQDISPVTQAAEEIEL
                HPLVAAENDTFDIYAEPFDPIPDPVQHSVTQSYLTSTPNTLSQSWGNTTVPLSIPSDW
                FVQSGPDITFPTASMGTPFSPVTPALPTGPVFITGSDFYLHPTWYFARRRRKRIPLFF
                TDVAA"
polyA_signal    4545. .4551
                /gene="L2"
                /note="putative"
gene            5771. .7276
                /gene="L1"
CDS             5771. .7276
                /gene="L1"
                /note="5771 is position of first start codon in ORF L1;
                putative"
                /codon_start=1
                /product="major capsid protein"
                /protein_id="AAA46935.1"
                /db_xref="GI:496201"
                /translation="MWRPSDSTVYVPPPNPVSKVVATDAYVKRTNIFYHASSSRLLAV
                GHPYYSIKKVNKTVVPKVSGYQYRVFKVVLPDPNKFALPDSSLFDPTTQRLVWACTGL
                EVGRGQPLGVGVSGHPLLNKYDDVENSGGYGGNPGQDNRVNVGMDYKQTQLCMVGCAP
                PLGEHWGKGTQCSNTSVQNGDCPPLELITSVIQDGDMVDTGFGAMNFADLQTNKSDVP
                LDICGTVCKYPDYLQMAADPYGDRLFFYLRKEQMFARHFFNRAGTVGEPVPDDLLVKG
                GNNRSSVASSIYVHTPSGSLVSSEAQLFNKPYWLQKAQGHNNGICWGNHLFVTVVDTT
                RSTNMTLCASVSKSATYTNSDYKEYMRHVEEFDLQFIFQLCSITLSAEVMAYIHTMNP
                SVLEDWNFGLSPPPNGTLEDTYRYVQSQAITCQKPTPEKEKQDPYKDMSFWEVNLKEK
                FSSELDQFPLGRKFLLQSGYRGRTSARTGIKRPAVSKPSTAPKRKRTKTKK"
      5'UTR     join(7277. .7931,1. .101)
                /standard_name="LCR"
                /function="regulatory region"
      repeat_region 7320. .7360
                /standard_name="direct repeats"
                /note="putative"
      repeat_region 7339. .7374
                /standard_name="direct repeats"
                /note="putative"
      repeat_region 7374. .7403
                /note="putative"
      polyA_signal  7457. .7462
                /note="putative"
      protein_bind  7592. .7603
                /note="putative"
                /function="gene transcription"
                /bound_moiety="E2"
      polyA_signal  7748. .7753
                /note="putative"
      protein_bind  7890. .7901
                /note="putative"
                /function="gene transcription"
                /bound_moiety="E2"
BASE COUNT     2406 a   1519 c   1736 g   2270 t
ORIGIN      4557 bp upstream of HindIII site.

 PPH11  Length: 7931  November 28, 2001 14:10  Type: N  Check: 3689  ..
Found using 'papiss' (pappu445.key)

...

                                                                |------|
669     CTTTAACACAACATTACCAAATACTGACCTGTTGCTGTGGATGTGACAGCAACGTCCGAC
                                                              719    726

729     TGGTTGTGGAGTGCACAGACGGAGACATCAGACAACTACAAGACCTTT

...

                        -- Search Statistics --

Times:                      CPU                 Total Elapsed
                         00:00:00.03            00:00:01.00

Number of sequences searched:                       6
Number of sequence hits:                            1
Number of separate matches:                         1
Number of sequence hits saved:                      0
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: November 29, 2001, 14:47:07 ; Search time 1391.6 Seconds
(without alignments)
260.806 Million cell updates/sec

Title: SEQ1
Perfect score: 22
Sequence: 1 TGACTGTGAACGTTCGAGATGA 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 661134

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vi:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rod:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rod:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

%

Result No.		Score	Query Match	Length	DB	ID	Description
1		22	100.0	22	6	AX036945	AX036945 Sequence
2		22	100.0	22	6	AX046993	AX046993 Sequence
3		22	100.0	22	6	AX083675	AX083675 Sequence
4		22	100.0	22	6	AX135650	AX135650 Sequence
5		22	100.0	22	6	AX148636	AX148636 Sequence
6		21	95.5	22	6	AX083681	AX083681 Sequence
7		21	95.5	22	6	AX148642	AX148642 Sequence
8		20.4	92.7	22	6	AR148608	AR148608 Sequence
9		20.4	92.7	22	6	AX036946	AX036946 Sequence
10		20.4	92.7	22	6	AX083676	AX083676 Sequence
11		20.4	92.7	22	6	AX083678	AX083678 Sequence
12		20.4	92.7	22	6	AX148637	AX148637 Sequence
13		20.4	92.7	22	6	AX148639	AX148639 Sequence
14		20.2	91.8	22	6	AX148643	AX148643 Sequence
15		20	90.9	22	6	AX083682	AX083682 Sequence
16		20	90.9	22	6	AX174913	AX174913 Sequence
17		19.4	88.2	22	6	AX083680	AX083680 Sequence
18		19.4	88.2	22	6	AX148641	AX148641 Sequence
19		18.8	85.5	22	6	AR148607	AR148607 Sequence
20		18.8	85.5	22	6	AR148609	AR148609 Sequence
21		18.8	85.5	22	6	AR148616	AR148616 Sequence
22		18.8	85.5	22	6	AX036944	AX036944 Sequence
23		18.8	85.5	22	6	AX036952	AX036952 Sequence
24		18.8	85.5	22	6	AX135651	AX135651 Sequence
25		18.8	85.5	22	6	AX148644	AX148644 Sequence
26		18.8	85.5	22	6	AX148645	AX148645 Sequence
27		17.2	78.2	22	6	AR148610	AR148610 Sequence
28		17.2	78.2	22	6	AX135652	AX135652 Sequence
29		15.6	70.9	22	6	AR148611	AR148611 Sequence
30		15.6	70.9	22	6	AR148613	AR148613 Sequence
31		15.6	70.9	22	6	AR148614	AR148614 Sequence
32		15.6	70.9	93	4	EMA270463	AJ270463 Elephas m
c 33		15	68.2	23	6	AX083677	AX083677 Sequence
c 34		15	68.2	23	6	AX148638	AX148638 Sequence
35		14.6	66.4	72	11	G42179	G42179 MMS09 Human
36		14	63.6	77	6	I40727	I40727 Sequence 58
37		14	63.6	93	4	MLA270467	AJ270467 Micropota
c 38		13.6	61.8	62	6	AX011500	AX011500 Sequence
39		13.6	61.8	77	6	AR125945	AR125945 Sequence
40		13.6	61.8	77	6	I47265	I47265 Sequence 19
41		13.6	61.8	93	9	HSPA5B8	Z79354 H.sapiens f
c 42		13.6	61.8	97	6	I35460	I35460 Sequence 11
c 43		13.6	61.8	98	6	I35457	I35457 Sequence 8
c 44		13.6	61.8	98	6	I35468	I35468 Sequence 19
c 45		13.6	61.8	98	6	I35471	I35471 Sequence 22

ALIGNMENTS

RESULT 1
AX036945
LOCUS AX036945 22 bp DNA PAT 16-NOV-2000
DEFINITION Sequence 2 from Patent FR2790955.
ACCESSION AX036945
VERSION AX036945.1 GI:11226373
KEYWORDS .
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 22)
AUTHORS Carpentier,A.
JOURNAL Patent: FR 2790955-A 2 22-SEP-2000;
ASSIST PUBL HOPITAUX DE PARIS (FR)
FEATURES Location/Qualifiers
source 1. .22
/organism="synthetic construct"
/db_xref="taxon:32630"
/note=" oligodesoxynucleotide"
BASE COUNT 6 a 3 c 7 g 6 t

```
ORIGIN


  Query Match             100.0%;  Score 22;  DB 6;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 0.13;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          ||||||||||||||||||||||
Db      1 TGACTGTGAACGTTCGAGATGA 22


RESULT   2
AX046993
LOCUS       AX046993       22 bp    DNA             PAT       15-DEC-2000
DEFINITION  Sequence 2 from Patent WO0067787.
ACCESSION   AX046993
VERSION     AX046993.1  GI:11876420
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Moss,R.B.
  TITLE     Hiv immunogenic compositions and methods
  JOURNAL   Patent: WO 0067787-A 2 16-NOV-2000;
            THE IMMUNE RESPONSE CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="phosphorothioate-modified synthetic
                     oligodeoxynucleotide"
BASE COUNT        6 a      3 c      7 g      6 t
ORIGIN


  Query Match             100.0%;  Score 22;  DB 6;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 0.13;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          ||||||||||||||||||||||
Db      1 TGACTGTGAACGTTCGAGATGA 22


RESULT   3
AX083675
LOCUS       AX083675       22 bp    DNA             PAT       28-FEB-2001
DEFINITION  Sequence 1 from Patent WO0112223.
ACCESSION   AX083675
VERSION     AX083675.1  GI:13185407
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   van Nest,G.
  TITLE     Methods of modulating an immune response using immunostimulatory s
            equences and compositions for use therein
  JOURNAL   Patent: WO 0112223-A 1 22-FEB-2001;
            Dynavax Technologies Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Synthetic construct"
BASE COUNT        6 a      3 c      7 g      6 t
ORIGIN


  Query Match             100.0%;  Score 22;  DB 6;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 0.13;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          ||||||||||||||||||||||
Db      1 TGACTGTGAACGTTCGAGATGA 22


RESULT   4
AX135650
LOCUS       AX135650       22 bp    DNA             PAT       29-MAY-2001
DEFINITION  Sequence 21 from Patent WO0132877.
ACCESSION   AX135650
VERSION     AX135650.1  GI:14271920
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Mackichan,M.L.
  TITLE     Cpg receptor (cpg-r) and methods relating thereto
  JOURNAL   Patent: WO 0132877-A 21 10-MAY-2001;
            CHIRON CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="CpG oligonucleotide"
BASE COUNT        6 a      3 c      7 g      6 t
ORIGIN


  Query Match             100.0%;  Score 22;  DB 6;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 0.13;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          ||||||||||||||||||||||
Db      1 TGACTGTGAACGTTCGAGATGA 22


RESULT   5
AX148636
LOCUS       AX148636       22 bp    DNA             PAT       08-JUN-2001
DEFINITION  Sequence 1 from Patent WO0135991.
ACCESSION   AX148636
VERSION     AX148636.1  GI:14347254
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Tuck,S. and van Nest,G.
  TITLE     Immunomodulatory compositions containing an immunostimulatory
            sequence linked to antigen and methods of use thereof
  JOURNAL   Patent: WO 0135991-A 1 25-MAY-2001;
            Dynavax Technologies Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic construct"
BASE COUNT        6 a      3 c      7 g      6 t
ORIGIN


  Query Match             100.0%;  Score 22;  DB 6;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 0.13;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          ||||||||||||||||||||||
Db      1 TGACTGTGAACGTTCGAGATGA 22
```

```
RESULT   6
AX083681
LOCUS       AX083681       22 bp    DNA             PAT       28-FEB-2001
DEFINITION  Sequence 7 from Patent WO0112223.
ACCESSION   AX083681
VERSION     AX083681.1  GI:13185413
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   van Nest,G.
  TITLE     Methods of modulating an immune response using immunostimulatory s
            equences and compositions for use therein
  JOURNAL   Patent: WO 0112223-A 7 22-FEB-2001;
            Dynavax Technologies Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     modified_base   11
                     /note="5-bromocytosine"
                     /mod_base=OTHER
BASE COUNT        6 a      2 c      7 g      6 t      1 others
ORIGIN


  Query Match             95.5%;  Score 21;  DB 6;  Length 22;
  Best Local Similarity   95.5%;  Pred. No. 0.49;
  Matches   21;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          |||||||||| ||||||||||
Db      1 TGACTGTGAANGTTCGAGATGA 22


RESULT   7
AX148642
LOCUS       AX148642       22 bp    DNA             PAT       08-JUN-2001
DEFINITION  Sequence 7 from Patent WO0135991.
ACCESSION   AX148642
VERSION     AX148642.1  GI:14347260
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Tuck,S. and van Nest,G.
  TITLE     Immunomodulatory compositions containing an immunostimulatory
            sequence linked to antigen and methods of use thereof
  JOURNAL   Patent: WO 0135991-A 7 25-MAY-2001;
            Dynavax Technologies Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic construct"
     modified_base   11
                     /note="5-bromocytosine"
                     /mod_base=OTHER
BASE COUNT        6 a      2 c      7 g      6 t      1 others
ORIGIN


  Query Match             95.5%;  Score 21;  DB 6;  Length 22;
  Best Local Similarity   95.5%;  Pred. No. 0.49;
  Matches   21;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          |||||||||| ||||||||||
Db      1 TGACTGTGAANGTTCGAGATGA 22


RESULT   8
AR148608
LOCUS       AR148608       22 bp    DNA             PAT       08-AUG-2001
DEFINITION  Sequence 2 from patent US 6225292.
ACCESSION   AR148608
VERSION     AR148608.1  GI:15112698
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Raz,E. and Roman,M.
  TITLE     Inhibitors of DNA immunostimulatory sequence activity
  JOURNAL   Patent: US 6225292-A 2 01-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="unknown"
BASE COUNT        7 a      2 c      7 g      6 t
ORIGIN


  Query Match             92.7%;  Score 20.4;  DB 6;  Length 22;
  Best Local Similarity   95.5%;  Pred. No. 1.1;
  Matches   21;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          |||||||||||||| |||||||
Db      1 TGACTGTGAACGTTAGAGATGA 22


RESULT   9
AX036946
LOCUS       AX036946       22 bp    DNA             PAT       16-NOV-2000
DEFINITION  Sequence 3 from Patent FR2790955.
ACCESSION   AX036946
VERSION     AX036946.1  GI:11226374
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Carpentier,A.
  JOURNAL   Patent: FR 2790955-A 3 22-SEP-2000;
            ASSIST PUBL HOPITAUX DE PARIS (FR)
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note=" oligodesoxynucleotide"
BASE COUNT        6 a      4 c      6 g      6 t
ORIGIN


  Query Match             92.7%;  Score 20.4;  DB 6;  Length 22;
  Best Local Similarity   95.5%;  Pred. No. 1.1;
  Matches   21;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          |||||||||||||| |||||||
Db      1 TGACTGTGAACGTTCCAGATGA 22


RESULT  10
AX083676
LOCUS       AX083676       22 bp    DNA             PAT       28-FEB-2001
DEFINITION  Sequence 2 from Patent WO0112223.
ACCESSION   AX083676
VERSION     AX083676.1  GI:13185408
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
```

```
REFERENCE   1  (bases 1 to 22)
  AUTHORS   van Nest,G.
  TITLE     Methods of modulating an immune response using immunostimulatory s
            equences and compositions for use therein
  JOURNAL   Patent: WO 0112223-A 2 22-FEB-2001;
            Dynavax Technologies Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Synthetic construct"
BASE COUNT        6 a      4 c      7 g      5 t
ORIGIN


  Query Match           92.7%;  Score 20.4;  DB 6;  Length 22;
  Best Local Similarity  95.5%;  Pred. No. 1.1;
  Matches   21;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          |||| ||||||||||||||||||
Db      1 TGACCGTGAACGTTCGAGATGA 22


RESULT  11
AX083678
LOCUS       AX083678       22 bp    DNA             PAT       28-FEB-2001
DEFINITION  Sequence 4 from Patent WO0112223.
ACCESSION   AX083678
VERSION     AX083678.1  GI:13185410
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   van Nest,G.
  TITLE     Methods of modulating an immune response using immunostimulatory s
            equences and compositions for use therein
  JOURNAL   Patent: WO 0112223-A 4 22-FEB-2001;
            Dynavax Technologies Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Synthetic construct"
BASE COUNT        6 a      4 c      6 g      6 t
ORIGIN


  Query Match           92.7%;  Score 20.4;  DB 6;  Length 22;
  Best Local Similarity  95.5%;  Pred. No. 1.1;
  Matches   21;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          |||||||||||||| ||||||
Db      1 TGACTGTGAACGTTCCAGATGA 22


RESULT  12
AX148637
LOCUS       AX148637       22 bp    DNA             PAT       08-JUN-2001
DEFINITION  Sequence 2 from Patent WO0135991.
ACCESSION   AX148637
VERSION     AX148637.1  GI:14347255
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Tuck,S. and van Nest,G.
  TITLE     Immunomodulatory compositions containing an immunostimulatory
            sequence linked to antigen and methods of use thereof
  JOURNAL   Patent: WO 0135991-A 2 25-MAY-2001;
            Dynavax Technologies Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic construct"
BASE COUNT        6 a      4 c      7 g      5 t
ORIGIN


  Query Match           92.7%;  Score 20.4;  DB 6;  Length 22;
  Best Local Similarity  95.5%;  Pred. No. 1.1;
  Matches   21;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          |||| ||||||||||||||||||
Db      1 TGACCGTGAACGTTCGAGATGA 22


RESULT  13
AX148639
LOCUS       AX148639       22 bp    DNA             PAT       08-JUN-2001
DEFINITION  Sequence 4 from Patent WO0135991.
ACCESSION   AX148639
VERSION     AX148639.1  GI:14347257
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Tuck,S. and van Nest,G.
  TITLE     Immunomodulatory compositions containing an immunostimulatory
            sequence linked to antigen and methods of use thereof
  JOURNAL   Patent: WO 0135991-A 4 25-MAY-2001;
            Dynavax Technologies Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic construct"
BASE COUNT        6 a      4 c      6 g      6 t
ORIGIN


  Query Match           92.7%;  Score 20.4;  DB 6;  Length 22;
  Best Local Similarity  95.5%;  Pred. No. 1.1;
  Matches   21;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          |||||||||||||| ||||||
Db      1 TGACTGTGAACGTTCCAGATGA 22


RESULT  14
AX148643
LOCUS       AX148643       22 bp    DNA             PAT       08-JUN-2001
DEFINITION  Sequence 8 from Patent WO0135991.
ACCESSION   AX148643
VERSION     AX148643.1  GI:14347261
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Tuck,S. and van Nest,G.
  TITLE     Immunomodulatory compositions containing an immunostimulatory
            sequence linked to antigen and methods of use thereof
  JOURNAL   Patent: WO 0135991-A 8 25-MAY-2001;
            Dynavax Technologies Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
```

```
                     /note="synthetic construct"
     modified_base   11
                     /note="5-bromocytosine"
                     /mod_base=OTHER
BASE COUNT        6 a      1 c      7 g      6 t      2 others
ORIGIN


  Query Match             91.8%;  Score 20.2;  DB 6;  Length 22;
  Best Local Similarity   90.9%;  Pred. No. 1.4;
  Matches   20;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          |||||||||| |||:|||||||
Db      1 TGACTGTGAANGTTBGAGATGA 22


RESULT  15
AX083682
LOCUS       AX083682       22 bp    DNA             PAT       28-FEB-2001
DEFINITION  Sequence 8 from Patent WO0112223.
ACCESSION   AX083682
VERSION     AX083682.1  GI:13185414
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   van Nest,G.
  TITLE     Methods of modulating an immune response using immunostimulatory s
            equences and compositions for use therein
  JOURNAL   Patent: WO 0112223-A 8 22-FEB-2001;
            Dynavax Technologies Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     modified_base   11
                     /note="5-bromocytosine"
                     /mod_base=OTHER
     modified_base   15
                     /note="5-bromocytosine"
                     /mod_base=OTHER
BASE COUNT        6 a      1 c      7 g      6 t      2 others
ORIGIN


  Query Match             90.9%;  Score 20;  DB 6;  Length 22;
  Best Local Similarity   90.9%;  Pred. No. 1.8;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          |||||||||| ||| |||||||
Db      1 TGACTGTGAANGTTNGAGATGA 22


RESULT  16
AX174913
LOCUS       AX174913       22 bp    DNA             PAT       03-JUL-2001
DEFINITION  Sequence 1 from Patent WO0143778.
ACCESSION   AX174913
VERSION     AX174913.1  GI:14598409
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Felgner,P.L. and Zelphati,O.
  TITLE     Use of cationic lipids for intracellular protein delivery
  JOURNAL   Patent: WO 0143778-A 1 21-JUN-2001;
            Gene Therapy Systems, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Synthetic peptide"
     modified_base   1
                     /note="n=T-NH2"
                     /mod_base=OTHER
     modified_base   22
                     /note="n=A-Rhodamine"
                     /mod_base=OTHER
BASE COUNT        5 a      3 c      7 g      5 t      2 others
ORIGIN


  Query Match             90.9%;  Score 20;  DB 6;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 1.8;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 GACTGTGAACGTTCGAGATG 21
          ||||||||||||||||||||
Db      2 GACTGTGAACGTTCGAGATG 21


RESULT  17
AX083680
LOCUS       AX083680       22 bp    DNA             PAT       28-FEB-2001
DEFINITION  Sequence 6 from Patent WO0112223.
ACCESSION   AX083680
VERSION     AX083680.1  GI:13185412
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   van Nest,G.
  TITLE     Methods of modulating an immune response using immunostimulatory s
            equences and compositions for use therein
  JOURNAL   Patent: WO 0112223-A 6 22-FEB-2001;
            Dynavax Technologies Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     modified_base   11
                     /note="5-bromocytosine"
                     /mod_base=OTHER
BASE COUNT        6 a      3 c      6 g      6 t      1 others
ORIGIN


  Query Match             88.2%;  Score 19.4;  DB 6;  Length 22;
  Best Local Similarity   90.9%;  Pred. No. 3.9;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          |||||||||| |||| ||||||
Db      1 TGACTGTGAANGTTCCAGATGA 22


RESULT  18
AX148641
LOCUS       AX148641       22 bp    DNA             PAT       08-JUN-2001
DEFINITION  Sequence 6 from Patent WO0135991.
ACCESSION   AX148641
VERSION     AX148641.1  GI:14347259
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Tuck,S. and van Nest,G.
  TITLE     Immunomodulatory compositions containing an immunostimulatory
            sequence linked to antigen and methods of use thereof
  JOURNAL   Patent: WO 0135991-A 6 25-MAY-2001;
```

```
           Dynavax Technologies Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic construct"
     modified_base   11
                     /note="5-bromocytosine"
                     /mod_base=OTHER
BASE COUNT        6 a      3 c      6 g      6 t      1 others
ORIGIN


  Query Match             88.2%;  Score 19.4;  DB 6;  Length 22;
  Best Local Similarity   90.9%;  Pred. No. 3.9;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          ||||||||||  |||| ||||||
Db      1 TGACTGTGAANGTTCCAGATGA 22


RESULT  19
AR148607
LOCUS       AR148607       22 bp    DNA             PAT       08-AUG-2001
DEFINITION  Sequence 1 from patent US 6225292.
ACCESSION   AR148607
VERSION     AR148607.1  GI:15112697
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Raz,E. and Roman,M.
  TITLE     Inhibitors of DNA immunostimulatory sequence activity
  JOURNAL   Patent: US 6225292-A 1 01-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="unknown"
BASE COUNT        7 a      1 c      8 g      6 t
ORIGIN


  Query Match             85.5%;  Score 18.8;  DB 6;  Length 22;
  Best Local Similarity   90.9%;  Pred. No. 8.6;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          ||||||||||  ||| |||||||
Db      1 TGACTGTGAAGGTTAGAGATGA 22


RESULT  20
AR148609
LOCUS       AR148609       22 bp    DNA             PAT       08-AUG-2001
DEFINITION  Sequence 3 from patent US 6225292.
ACCESSION   AR148609
VERSION     AR148609.1  GI:15112699
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Raz,E. and Roman,M.
  TITLE     Inhibitors of DNA immunostimulatory sequence activity
  JOURNAL   Patent: US 6225292-A 3 01-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="unknown"
BASE COUNT        7 a      3 c      6 g      6 t
ORIGIN


  Query Match             85.5%;  Score 18.8;  DB 6;  Length 22;
  Best Local Similarity   90.9%;  Pred. No. 8.6;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          ||||||||||  || |||||||
Db      1 TGACTGTGAACCTTAGAGATGA 22


RESULT  21
AR148616
LOCUS       AR148616       22 bp    DNA             PAT       08-AUG-2001
DEFINITION  Sequence 10 from patent US 6225292.
ACCESSION   AR148616
VERSION     AR148616.1  GI:15112706
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Raz,E. and Roman,M.
  TITLE     Inhibitors of DNA immunostimulatory sequence activity
  JOURNAL   Patent: US 6225292-A 10 01-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="unknown"
BASE COUNT        7 a      1 c      7 g      7 t
ORIGIN


  Query Match             85.5%;  Score 18.8;  DB 6;  Length 22;
  Best Local Similarity   90.9%;  Pred. No. 8.6;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          ||||||||||  ||| |||||||
Db      1 TGACTGTGAATGTTAGAGATGA 22


RESULT  22
AX036944
LOCUS       AX036944       22 bp    DNA             PAT       16-NOV-2000
DEFINITION  Sequence 1 from Patent FR2790955.
ACCESSION   AX036944
VERSION     AX036944.1  GI:11226372
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Carpentier,A.
  JOURNAL   Patent: FR 2790955-A 1 22-SEP-2000;
            ASSIST PUBL HOPITAUX DE PARIS (FR)
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note=" oligodesoxynucleotide"
BASE COUNT        7 a      1 c      8 g      6 t
ORIGIN


  Query Match             85.5%;  Score 18.8;  DB 6;  Length 22;
  Best Local Similarity   90.9%;  Pred. No. 8.6;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          ||||||||||  ||| |||||||
Db      1 TGACTGTGAAGGTTAGAGATGA 22


RESULT  23
AX036952
```

```
LOCUS       AX036952       22 bp    DNA             PAT       16-NOV-2000
DEFINITION  Sequence 9 from Patent FR2790955.
ACCESSION   AX036952
VERSION     AX036952.1  GI:11226380
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Carpentier,A.
  JOURNAL   Patent: FR 2790955-A 9 22-SEP-2000;
            ASSIST PUBL HOPITAUX DE PARIS (FR)
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note=" oligodesoxynucleotide"
BASE COUNT        7 a      2 c      6 g      7 t
ORIGIN


  Query Match             85.5%;  Score 18.8;  DB 6;  Length 22;
  Best Local Similarity   90.9%;  Pred. No. 8.6;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          ||||||||||||||  ||||||
Db      1 TGACTGTGAACGTTATAGATGA 22


RESULT  24
AX135651
LOCUS       AX135651       22 bp    DNA             PAT       29-MAY-2001
DEFINITION  Sequence 22 from Patent WO0132877.
ACCESSION   AX135651
VERSION     AX135651.1  GI:14271921
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Mackichan,M.L.
  TITLE     Cpg receptor (cpg-r) and methods relating thereto
  JOURNAL   Patent: WO 0132877-A 22 10-MAY-2001;
            CHIRON CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="CpG oligonucleotide"
BASE COUNT        6 a      3 c      7 g      6 t
ORIGIN


  Query Match             85.5%;  Score 18.8;  DB 6;  Length 22;
  Best Local Similarity   90.9%;  Pred. No. 8.6;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          ||||||||||||||  | |||||
Db      1 TGACTGTGAACGTTAGCGATGA 22


RESULT  25
AX148644
LOCUS       AX148644       22 bp    DNA             PAT       08-JUN-2001
DEFINITION  Sequence 9 from Patent WO0135991.
ACCESSION   AX148644
VERSION     AX148644.1  GI:14347262
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Tuck,S. and van Nest,G.
  TITLE     Immunomodulatory compositions containing an immunostimulatory
            sequence linked to antigen and methods of use thereof
  JOURNAL   Patent: WO 0135991-A 9 25-MAY-2001;
            Dynavax Technologies Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic construct"
BASE COUNT        7 a      1 c      8 g      6 t
ORIGIN


  Query Match             85.5%;  Score 18.8;  DB 6;  Length 22;
  Best Local Similarity   90.9%;  Pred. No. 8.6;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          ||||||||||  ||| |||||||
Db      1 TGACTGTGAAGGTTAGAGATGA 22


RESULT  26
AX148645
LOCUS       AX148645       22 bp    DNA             PAT       08-JUN-2001
DEFINITION  Sequence 10 from Patent WO0135991.
ACCESSION   AX148645
VERSION     AX148645.1  GI:14347263
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Tuck,S. and van Nest,G.
  TITLE     Immunomodulatory compositions containing an immunostimulatory
            sequence linked to antigen and methods of use thereof
  JOURNAL   Patent: WO 0135991-A 10 25-MAY-2001;
            Dynavax Technologies Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic construct"
BASE COUNT        7 a      3 c      6 g      6 t
ORIGIN


  Query Match             85.5%;  Score 18.8;  DB 6;  Length 22;
  Best Local Similarity   90.9%;  Pred. No. 8.6;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          ||||||||||| || |||||||
Db      1 TGACTGTGAACCTTAGAGATGA 22


RESULT  27
AR148610
LOCUS       AR148610       22 bp    DNA             PAT       08-AUG-2001
DEFINITION  Sequence 4 from patent US 6225292.
ACCESSION   AR148610
VERSION     AR148610.1  GI:15112700
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Raz,E. and Roman,M.
  TITLE     Inhibitors of DNA immunostimulatory sequence activity
  JOURNAL   Patent: US 6225292-A 4 01-MAY-2001;
FEATURES             Location/Qualifiers
```

```
     source          1. .22
                     /organism="unknown"
BASE COUNT        7 a      2 c      7 g      6 t
ORIGIN


  Query Match             78.2%;  Score 17.2;  DB 6;  Length 22;
  Best Local Similarity   86.4%;  Pred. No. 69;
  Matches   19;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          ||||||||||  || |||||||
Db      1 TGACTGTGAAGCTTAGAGATGA 22


RESULT  28
AX135652
LOCUS       AX135652       22 bp    DNA             PAT       29-MAY-2001
DEFINITION  Sequence 23 from Patent WO0132877.
ACCESSION   AX135652
VERSION     AX135652.1  GI:14271922
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Mackichan,M.L.
  TITLE     Cpg receptor (cpg-r) and methods relating thereto
  JOURNAL   Patent: WO 0132877-A 23 10-MAY-2001;
            CHIRON CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="CpG oligonucleotide"
BASE COUNT        6 a      3 c      8 g      5 t
ORIGIN


  Query Match             78.2%;  Score 17.2;  DB 6;  Length 22;
  Best Local Similarity   86.4%;  Pred. No. 69;
  Matches   19;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          ||||||||||||||| |||  ||
Db      1 TGACTGTGAACGTTAGAGCGGA 22


RESULT  29
AR148611
LOCUS       AR148611       22 bp    DNA             PAT       08-AUG-2001
DEFINITION  Sequence 5 from patent US 6225292.
ACCESSION   AR148611
VERSION     AR148611.1  GI:15112701
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Raz,E. and Roman,M.
  TITLE     Inhibitors of DNA immunostimulatory sequence activity
  JOURNAL   Patent: US 6225292-A 5 01-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="unknown"
BASE COUNT        5 a      3 c      6 g      8 t
ORIGIN


  Query Match             70.9%;  Score 15.6;  DB 6;  Length 22;
  Best Local Similarity   81.8%;  Pred. No. 5.6e+02;
  Matches   18;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          ||||||||  | || |||||||
Db      1 TGACTGTGTTCCTTAGAGATGA 22


RESULT  30
AR148613
LOCUS       AR148613       22 bp    DNA             PAT       08-AUG-2001
DEFINITION  Sequence 7 from patent US 6225292.
ACCESSION   AR148613
VERSION     AR148613.1  GI:15112703
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Raz,E. and Roman,M.
  TITLE     Inhibitors of DNA immunostimulatory sequence activity
  JOURNAL   Patent: US 6225292-A 7 01-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="unknown"
BASE COUNT        6 a      2 c      9 g      5 t
ORIGIN


  Query Match             70.9%;  Score 15.6;  DB 6;  Length 22;
  Best Local Similarity   81.8%;  Pred. No. 5.6e+02;
  Matches   18;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          ||||||||  | ||  |||||||
Db      1 TGACTGTGGAGGTCAGAGATGA 22


RESULT  31
AR148614
LOCUS       AR148614       22 bp    DNA             PAT       08-AUG-2001
DEFINITION  Sequence 8 from patent US 6225292.
ACCESSION   AR148614
VERSION     AR148614.1  GI:15112704
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Raz,E. and Roman,M.
  TITLE     Inhibitors of DNA immunostimulatory sequence activity
  JOURNAL   Patent: US 6225292-A 8 01-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="unknown"
BASE COUNT        6 a      2 c      9 g      5 t
ORIGIN


  Query Match             70.9%;  Score 15.6;  DB 6;  Length 22;
  Best Local Similarity   81.8%;  Pred. No. 5.6e+02;
  Matches   18;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          |||||||||  | | |||||||
Db      1 TGACTGTGAGGGCTAGAGATGA 22


RESULT  32
EMA270463
LOCUS       EMA270463      93 bp    DNA             MAM       04-JAN-2001
DEFINITION  Elephas maximus partial cryaa gene for alpha-A crystallin chain,
            exon 2.
ACCESSION   AJ270463
VERSION     AJ270463.1  GI:10803352
KEYWORDS    alpha-A crystallin chain; cryaA gene.
```

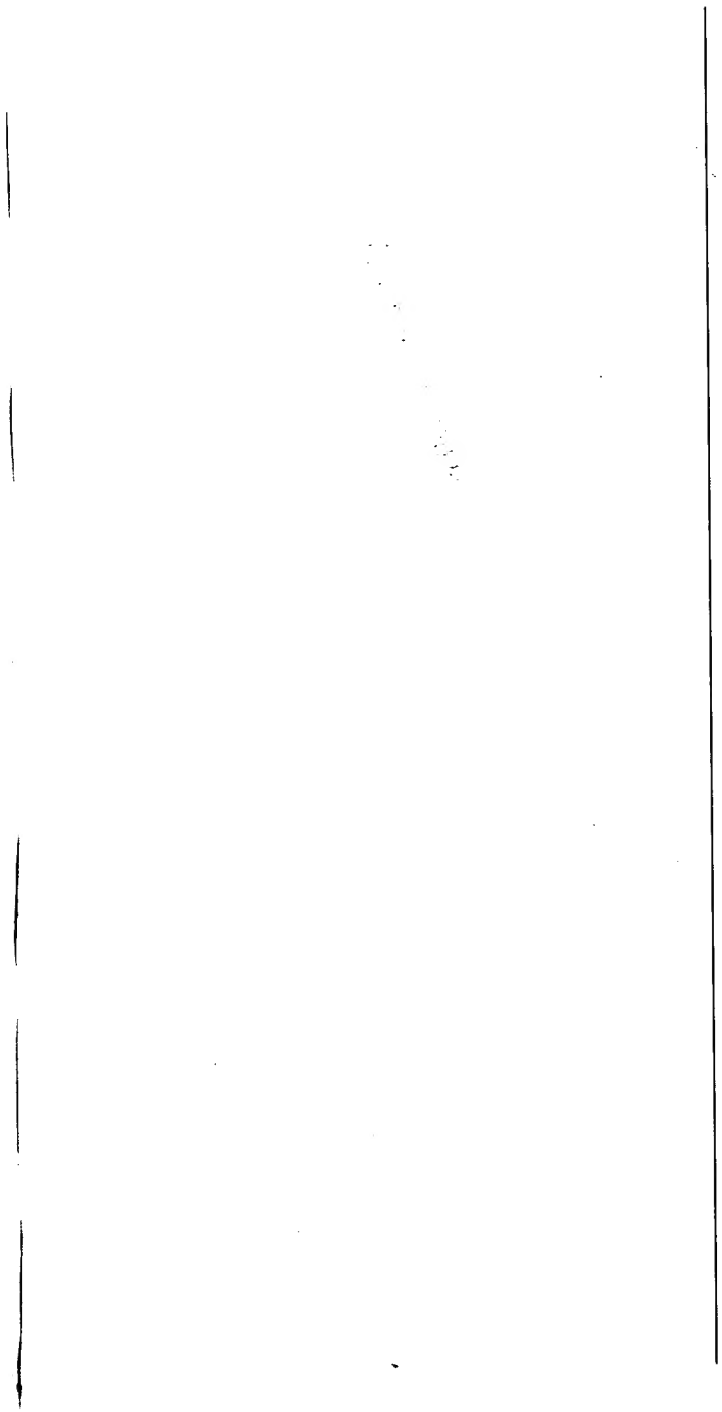
```
SOURCE      Asiatic elephant.
  ORGANISM  Elephas maximus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Proboscidea; Elephantidae; Elephas.
REFERENCE   1  (bases 1 to 93)
  AUTHORS   van Dijk,M.A., Madsen,O., Catzeflis,F., Stanhope,M.J., de Jong,W.W.
            and Pagel,M.
  TITLE     From the Cover: Protein sequence signatures support the African
            clade of mammals
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (1), 188-193 (2001)
   PUBMED   11114173
REFERENCE   2  (bases 1 to 93)
  AUTHORS   van Dijk,M.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-OCT-1999) van Dijk M.A., Department of Biochemistry,
            University of Nijmegen, P.O. Box 9101, 6500 HB, NETHERLANDS
FEATURES             Location/Qualifiers
     source          1. .93
                     /organism="Elephas maximus"
                     /db_xref="taxon:9783"
     exon            <1. .>93
                     /gene="cryaa"
                     /number=2
     gene            1. .93
                     /gene="cryaa"
     CDS             <1. .>93
                     /gene="cryaa"
                     /function="may contribute to the transparency and
                     refractive index of the lens"
                     /codon_start=1
                     /product="alpha-A crystallin chain"
                     /protein_id="CAC13113.1"
                     /db_xref="GI:10803353"
                     /translation="VRSDRDQFLILLDVKHFSPEDLTVKVQDDFV"
BASE COUNT       16 a      26 c      27 g      24 t
ORIGIN


  Query Match             70.9%;  Score 15.6;  DB 4;  Length 93;
  Best Local Similarity   81.8%;  Pred. No. 6.2e+02;
  Matches   18;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           |||||||||| || |  |||||
Db      65 TGACTGTGAAGGTGCAGGATGA 86


RESULT  33
AX083677/c
LOCUS       AX083677       23 bp    DNA             PAT       28-FEB-2001
DEFINITION  Sequence 3 from Patent WO0112223.
ACCESSION   AX083677
VERSION     AX083677.1  GI:13185409
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 23)
  AUTHORS   van Nest,G.
  TITLE     Methods of modulating an immune response using immunostimulatory s
            equences and compositions for use therein
  JOURNAL   Patent: WO 0112223-A 3 22-FEB-2001;
            Dynavax Technologies Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .23
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Synthetic construct"
BASE COUNT        6 a      8 c      3 g      6 t
ORIGIN


  Query Match             68.2%;  Score 15;  DB 6;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+03;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       8 GAACGTTCGAGATGA 22
           |||||||||||||||
Db      15 GAACGTTCGAGATGA 1


RESULT  34
AX148638/c
LOCUS       AX148638       23 bp    DNA             PAT       08-JUN-2001
DEFINITION  Sequence 3 from Patent WO0135991.
ACCESSION   AX148638
VERSION     AX148638.1  GI:14347256
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 23)
  AUTHORS   Tuck,S. and van Nest,G.
  TITLE     Immunomodulatory compositions containing an immunostimulatory
            sequence linked to antigen and methods of use thereof
  JOURNAL   Patent: WO 0135991-A 3 25-MAY-2001;
            Dynavax Technologies Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .23
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic construct"
BASE COUNT        6 a      8 c      3 g      6 t
ORIGIN


  Query Match             68.2%;  Score 15;  DB 6;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+03;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       8 GAACGTTCGAGATGA 22
           |||||||||||||||
Db      15 GAACGTTCGAGATGA 1


RESULT  35
G42179
LOCUS       G42179         72 bp    DNA             STS       17-JUN-1999
DEFINITION  MMS09 Human Homo sapiens STS genomic, sequence tagged site.
ACCESSION   G42179
VERSION     G42179.1  GI:4731081
KEYWORDS    STS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 72)
  AUTHORS   Sohocki,M.M., Malone,K.A., Sullivan,L.S. and Daiger,S.P.
  TITLE     Localization of retina/pineal-expressed sequences: identification
            of novel candidate genes for inherited retinal disorders
  JOURNAL   Genomics 58 (1), 29-33 (1999)
  MEDLINE   99265969
COMMENT
            Contact: Melanie M Sohocki
            University of Texas Health Science Center, Houston
            PO Box 20334, Houston, TX 77225-0334, USA
            Tel: 713-500-9841
            Fax: 713-500-0900
            Email: msohocki@gsbs3.gs.uth.tmc.edu
            Primer A: TGCTGACTGTGAACCTACGG
            Primer B: ATGCTAGGGCATCATCCTTG
            STS size: 72
            PCR Profile:
                 Presoak:                95 degrees C for 5.00 minute(s)
                 Denaturation          95 degrees C for 1.00 minute(s)
                 Annealing               58 degrees C for 1.00 minute(s)
```

```
                Polymerization       72 degrees C for 1.00 minute(s)
                PCR Cycles           40
                Thermal Cycler: MJ Research PTC-200
            Protocol:
                        Template:            50-150 ng
                        Primer:         each 10 uM
                        DNTPs:          each 200 uL
                        Taq Polymerase       0.05 units/ul
                        Total Vol:           20 ul

            Buffer:
                        MgCl2:               2.5 mM
                        KCl:                 50 mM
                        Tris-HCl:            10 mM
                        DMSO                 5%
                        pH:                  8.3.
FEATURES                Location/Qualifiers
     source             1. .72
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /map="10010000010000010101000000011000000100000100111000111
                        10000001000010202110021000002 0100100000001"
                        /clone_lib="Human"
     STS                1. .72
     primer_bind        1. .20
     primer_bind        complement(53. .72)
BASE COUNT       15 a      21 c      22 g      13 t       1 others
ORIGIN


  Query Match             66.4%;  Score 14.6;  DB 11;  Length 72;
  Best Local Similarity   81.0%;  Pred. No. 2.3e+03;
  Matches   17;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATG 21
           ||||||||||| | || || |
Db       4 TGACTGTGAACCTACGGGAAG 24


RESULT  36
I40727
LOCUS       I40727          77 bp     DNA             PAT       13-MAY-1997
DEFINITION  Sequence 58 from patent US 5622828.
ACCESSION   I40727
VERSION     I40727.1  GI:2082207
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 77)
  AUTHORS   Parma,D.H. and Gold,L.
  TITLE     High-affinity oligonucleotide ligands to secretory phospholipase A2
            (sPLA.sub.2)
  JOURNAL   Patent: US 5622828-A 58 22-APR-1997;
FEATURES                Location/Qualifiers
     source             1. .77
                        /organism="unknown"
BASE COUNT       22 a      26 c      19 g      10 t
ORIGIN


  Query Match             63.6%;  Score 14;  DB 6;  Length 77;
  Best Local Similarity   77.3%;  Pred. No. 5e+03;
  Matches   17;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           || |   |||||||||| ||||
Db      42 TGCCCACGAACGTTCGACATGA 63


RESULT  37
MLA270467
LOCUS       MLA270467       93 bp     DNA             MAM       04-JAN-2001
```

```
DEFINITION  Micropotamogale lamottei partial cryaa gene for alpha-A crystallin
            chain, exon 2.
ACCESSION   AJ270467
VERSION     AJ270467.1  GI:10803426
KEYWORDS    alpha-A crystallin chain; cryaA gene.
SOURCE      Nimba otter shrew.
  ORGANISM  Micropotamogale lamottei
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Insectivora; Tenrecidae; Micropotamogale.
REFERENCE   1  (bases 1 to 93)
  AUTHORS   van Dijk,M.A., Madsen,O., Catzeflis,F., Stanhope,M.J., de Jong,W.W.
            and Pagel,M.
  TITLE     From the Cover: Protein sequence signatures support the African
            clade of mammals
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (1), 188-193 (2001)
   PUBMED   11114173
REFERENCE   2  (bases 1 to 93)
  AUTHORS   van Dijk,M.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-OCT-1999) van Dijk M.A., Department of Biochemistry,
            University of Nijmegen, P.O. Box 9101, 6500 HB, NETHERLANDS
FEATURES                Location/Qualifiers
     source             1. .93
                        /organism="Micropotamogale lamottei"
                        /db_xref="taxon:105689"
                        /note="order tenrec"
     exon               <1. .>93
                        /gene="cryaa"
                        /number=2
     gene               1. .93
                        /gene="cryaa"
     CDS                <1. .>93
                        /gene="cryaa"
                        /function="may contribute to the transparency and
                        refractive index of the lens"
                        /codon_start=1
                        /product="alpha-A crystallin chain"
                        /protein_id="CAC13129.1"
                        /db_xref="GI:10803427"
                        /translation="VRSDRDQFLILLDVKHFSPEDLTVKVLEDFV"
BASE COUNT       16 a      26 c      30 g      21 t
ORIGIN


  Query Match             63.6%;  Score 14;  DB 4;  Length 93;
  Best Local Similarity   77.3%;  Pred. No. 5e+03;
  Matches   17;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           |||||||||| || |  || ||
Db      65 TGACTGTGAAGGTGCTGGAGGA 86


RESULT  38
AX011500/c
LOCUS       AX011500        62 bp     DNA             PAT       06-SEP-2000
DEFINITION  Sequence 177 from Patent WO9955907.
ACCESSION   AX011500
VERSION     AX011500.1  GI:9998050
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 62)
  AUTHORS   Koetter,P., Entian,K.D. and Diu-Hercend,A.
  TITLE     Method for screening antimycotic substances using essential genes
            from s. Cerevisiae
  JOURNAL   Patent: WO 9955907-A 177 04-NOV-1999;
            KOETTER PETER (DE); ENTIAN KARL DIETER (DE); DIU HERCEND ANITA
            (FR); HOECHST MARION ROUSSEL INC (FR)
FEATURES                Location/Qualifiers
     source             1. .62
                        /organism="synthetic construct"
```

```
                     /db_xref="taxon:32630"
                     /note="primer YDR472w-S2"
BASE COUNT       15 a      19 c      13 g      15 t
ORIGIN

  Query Match             61.8%;  Score 13.6;  DB 6;  Length 62;
  Best Local Similarity   80.0%;  Pred. No. 8.2e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGAT 20
           |||| | ||| ||| |||||
Db      26 TGACAGGGAAGGTTTGAGAT 7


RESULT  39
AR125945
LOCUS       AR125945        77 bp    DNA              PAT       16-MAY-2001
DEFINITION  Sequence 287 from patent US 6177557.
ACCESSION   AR125945
VERSION     AR125945.1  GI:14112007
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 77)
  AUTHORS   Janjic,N., Gold,L. and Tasset,D.
  TITLE     High affinity ligands of basic fibroblast growth factor and
            thrombin
  JOURNAL   Patent: US 6177557-A 287 23-JAN-2001;
FEATURES             Location/Qualifiers
     source          1. .77
                     /organism="unknown"
BASE COUNT       21 a      27 c      19 g      10 t
ORIGIN

  Query Match             61.8%;  Score 13.6;  DB 6;  Length 77;
  Best Local Similarity   80.0%;  Pred. No. 8.4e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       3 ACTGTGAACGTTCGAGATGA 22
           |||||| | ||||| ||||
Db      44 ACTGTGCCCCTTCGACATGA 63


RESULT  40
I47265
LOCUS       I47265          77 bp    DNA              PAT       07-OCT-1997
DEFINITION  Sequence 195 from patent US 5639868.
ACCESSION   I47265
VERSION     I47265.1  GI:2471230
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 77)
  AUTHORS   Janjic,N. and Gold,L.
  TITLE     High-affinity RNA ligands for basic fibroblast growth factor
  JOURNAL   Patent: US 5639868-A 195 17-JUN-1997;
FEATURES             Location/Qualifiers
     source          1. .77
                     /organism="unknown"
BASE COUNT       21 a      27 c      19 g      10 t
ORIGIN

  Query Match             61.8%;  Score 13.6;  DB 6;  Length 77;
  Best Local Similarity   80.0%;  Pred. No. 8.4e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       3 ACTGTGAACGTTCGAGATGA 22
           |||||| | ||||| ||||
Db      44 ACTGTGCCCCTTCGACATGA 63


RESULT  41
HSPA5B8
LOCUS       HSPA5B8         93 bp    DNA              PRI       23-AUG-1996
DEFINITION  H.sapiens flow-sorted chromosome 6 TaqI fragment, SC6pA5B8.
ACCESSION   Z79354
VERSION     Z79354.1  GI:1508632
KEYWORDS    Anonymous marker; single read.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 93)
  AUTHORS   Mungall,A.J., Huckle,E., Langford,C., Ross,M.T. and Rice,C.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-AUG-1996) The Sanger Centre, Wellcome Trust Genome
            Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail contact:
            humquery@sanger.ac.uk
COMMENT     Vector: pBSIISK+.
FEATURES             Location/Qualifiers
     source          1. .93
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /sex="female"
                     /dev_stage="adult"
                     /tissue_type="EBV lymphoblastoid cell line"
                     /clone_lib="SC6pA"
                     /clone="SC6pA5B8"
                     /note="The estimated purity of the flow-sorted chromosome
                     6 library is >97%"
BASE COUNT       33 a      17 c      21 g      22 t
ORIGIN

  Query Match             61.8%;  Score 13.6;  DB 9;  Length 93;
  Best Local Similarity   80.0%;  Pred. No. 8.5e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       3 ACTGTGAACGTTCGAGATGA 22
           | | || |||||| ||||||
Db      49 ATTTTGGACGTTCAAGATGA 68


RESULT  42
I35460/c
LOCUS       I35460          97 bp    DNA              PAT       13-MAY-1997
DEFINITION  Sequence 11 from patent US 5599917.
ACCESSION   I35460
VERSION     I35460.1  GI:2088428
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 97)
  AUTHORS   Coppola,G.R., Beutel,B.A. and Bertelsen,A.H.
  TITLE     Inhibition of interferon-.gamma. with oligonucleotides
  JOURNAL   Patent: US 5599917-A 11 04-FEB-1997;
FEATURES             Location/Qualifiers
     source          1. .97
                     /organism="unknown"
BASE COUNT       22 a      26 c      28 g      21 t
ORIGIN

  Query Match             61.8%;  Score 13.6;  DB 6;  Length 97;
  Best Local Similarity   80.0%;  Pred. No. 8.5e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       3 ACTGTGAACGTTCGAGATGA 22
           ||||||| |  |||||| ||
```

```
Db      97 ACTGTGACCTCTCGAGACGA 78


RESULT  43
I35457/c
LOCUS       I35457          98 bp    DNA             PAT       13-MAY-1997
DEFINITION  Sequence 8 from patent US 5599917.
ACCESSION   I35457
VERSION     I35457.1  GI:2088425
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 98)
  AUTHORS   Coppola,G.R., Beutel,B.A. and Bertelsen,A.H.
  TITLE     Inhibition of interferon-.gamma. with oligonucleotides
  JOURNAL   Patent: US 5599917-A 8 04-FEB-1997;
FEATURES             Location/Qualifiers
     source          1. .98
                     /organism="unknown"
BASE COUNT       25 a      16 c      29 g      28 t
ORIGIN


  Query Match             61.8%;  Score 13.6;  DB 6;  Length 98;
  Best Local Similarity   80.0%;  Pred. No. 8.5e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       3 ACTGTGAACGTTCGAGATGA 22
           ||||||| |  ||||| |||
Db      98 ACTGTGACCTCTCGAGGTGA 79


RESULT  44
I35468/c
LOCUS       I35468          98 bp    DNA             PAT       13-MAY-1997
DEFINITION  Sequence 19 from patent US 5599917.
ACCESSION   I35468
VERSION     I35468.1  GI:2088436
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 98)
  AUTHORS   Coppola,G.R., Beutel,B.A. and Bertelsen,A.H.
  TITLE     Inhibition of interferon-.gamma. with oligonucleotides
  JOURNAL   Patent: US 5599917-A 19 04-FEB-1997;
FEATURES             Location/Qualifiers
     source          1. .98
                     /organism="unknown"
BASE COUNT       23 a      17 c      28 g      30 t
ORIGIN


  Query Match             61.8%;  Score 13.6;  DB 6;  Length 98;
  Best Local Similarity   80.0%;  Pred. No. 8.5e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       3 ACTGTGAACGTTCGAGATGA 22
           ||||||| |  |||||| ||
Db      98 ACTGTGACCTCTCGAGACGA 79


RESULT  45
I35471/c
LOCUS       I35471          98 bp    DNA             PAT       13-MAY-1997
DEFINITION  Sequence 22 from patent US 5599917.
ACCESSION   I35471
VERSION     I35471.1  GI:2088439
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 98)
  AUTHORS   Coppola,G.R., Beutel,B.A. and Bertelsen,A.H.
  TITLE     Inhibition of interferon-.gamma. with oligonucleotides
  JOURNAL   Patent: US 5599917-A 22 04-FEB-1997;
FEATURES             Location/Qualifiers
     source          1. .98
                     /organism="unknown"
BASE COUNT       28 a      22 c      25 g      23 t
ORIGIN


  Query Match             61.8%;  Score 13.6;  DB 6;  Length 98;
  Best Local Similarity   80.0%;  Pred. No. 8.5e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       3 ACTGTGAACGTTCGAGATGA 22
           ||||||| |  ||||| |||
Db      98 ACTGTGACCTCTCGAGGTGA 79


Search completed: November 29, 2001, 14:47:08
Job time: 8321 sec
```



GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: November 29, 2001, 14:51:05 ; Search time 158.03 Seconds
(without alignments)
119.352 Million cell updates/sec

Title: SEQ1
Perfect score: 22
Sequence: 1 TGACTGTGAACGTTCGAGATGA 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1084414

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_1101:*
1: /SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	22	100.0	22	19	AAV32079	Nucleotide sequenc
2	22	100.0	22	20	AAX36624	ISS-ODN DY1018 nuc
3	22	100.0	22	20	AAV80097	Immunomodulatory o
4	22	100.0	22	20	AAV80102	Immunomodulatory o
5	22	100.0	22	20	AAV80103	Immunomodulatory o
6	22	100.0	22	21	AAC64051	Immunostimulatory
7	22	100.0	22	21	AAA96253	Sequence of a stab
8	22	100.0	22	21	AAA90458	CpG adjuvant oligo
9	22	100.0	22	21	AAA14467	Immunostimulatory
10	22	100.0	22	21	AAA38065	Immunostimulatory
11	22	100.0	22	21	AAA38071	Immunostimulatory
12	22	100.0	22	21	AAA38072	Immunostimulatory
13	22	100.0	22	21	AAZ55876	Immunomodulatory o
14	22	100.0	22	22	AAH42533	Phosphorothioate b
15	22	100.0	22	22	AAH73439	Immunomodulatory n
16	22	100.0	22	22	AAH44109	5' terminal NH2 gr
17	22	100.0	22	22	AAH41573	Immunostimulatory
18	22	100.0	22	22	AAH20403	CpG motif containi
19	22	100.0	22	22	AAF77040	Immunomodulatory D
20	22	100.0	22	22	AAF29800	Cholera toxin immu
21	22	100.0	22	22	AAC82107	Oligonucleotide OD
22	22	100.0	22	22	AAA92377	CG motif and CFA c
23	21.2	96.4	22	22	AAF77046	Immunostimulatory
24	21	95.5	22	21	AAZ55880	Immunomodulatory o
25	21	95.5	22	22	AAH41579	Immunostimulatory
c 26	20.4	92.7	22	20	AAV80105	Oligo used in expe
27	20.4	92.7	22	20	AAV80096	Immunomodulatory o
28	20.4	92.7	22	20	AAV80099	Immunomodulatory o
29	20.4	92.7	22	20	AAV80101	Immunomodulatory o
30	20.4	92.7	22	21	AAA96254	Sequence of a stab
31	20.4	92.7	22	21	AAA38066	Immunostimulatory
32	20.4	92.7	22	21	AAA38068	Immunostimulatory
33	20.4	92.7	22	21	AAA38070	Immunostimulatory
34	20.4	92.7	22	22	AAH42534	Phosphorothioate b
35	20.4	92.7	22	22	AAH73440	Immunomodulatory n
36	20.4	92.7	22	22	AAH41574	Immunostimulatory
37	20.4	92.7	22	22	AAH41576	Immunostimulatory
38	20.4	92.7	22	22	AAF77041	Immunostimulatory
39	20.4	92.7	22	22	AAF77043	Immunostimulatory
40	20.4	92.7	22	22	AAF77047	Immunostimulatory
41	20.2	91.8	22	22	AAH41580	Immunostimulatory
42	20	90.9	22	21	AAZ55881	Immunomodulatory o
43	19.6	89.1	22	22	AAF77045	Immunostimulatory
44	19.4	88.2	22	21	AAZ55877	Immunomodulatory o
45	19.4	88.2	22	22	AAH41578	Immunostimulatory

ALIGNMENTS

RESULT 1
AAV32079
ID AAV32079 standard; DNA; 22 BP.
XX
AC AAV32079;
XX
DT 09-SEP-1998 (first entry)
XX
DE Nucleotide sequence of DY1018.
XX
KW DY1018; beta-gal; ISS-PN/IMM; antigen; immune response; antibody;
KW immunisation; anaphylaxis; IgE; retinopathies; ss.
XX
OS synthetic.
XX
FH Key Location/Qualifiers
FT modified_base 1..22
FT /*tag= a
FT /note= "phosphothioate backbone"
XX
PN WO9816247-A1.
XX
PD 23-APR-1998.
XX
PF 09-OCT-1997; 97WO-US19004.
XX
PR 11-OCT-1996; 96US-0028118.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Carson DA, Raz E, Roman M;
XX
DR WPI; 1998-261028/23.
XX

```
PT   New immunomodulatory compositions - comprising an antigen conjugated
PT   to a polynucleotide that contains an immunostimulatory sequence
XX
PS   Example 1; Page 36; 69pp; English.
XX
CC   This is the nucleotide sequence of DY1018, which is conjugated to
CC   beta-gal to form ISS-PN/IMM, comprising an immunomodulatory molecule
CC   (IMM), which comprises an antigen conjugated to a polynucleotide
CC   (PN) that contains at least one immunostimulatory nucleotide sequence
CC   (ISS). The conjugate synergistically boost the magnitude of the host
CC   immune response against an antigen to a level greater that the host
CC   immune response to either the IMM, antigen or ISS-PN alone. These
CC   responses to ISS-PN/IMM conjugates are particularly acute during
CC   the important early phase of the host immune response to an antigen.
CC   The ISS-PN/IMM conjugates boost both humoral (antibody) and cellular
CC   (Th1 type) immune responses of the host. Thus, use of the method to
CC   boost the immune responsiveness of a host to subsequent challenge by a
CC   sensitising antigen without immunisation avoids the risk of
CC   Th2-mediated, immunisation-induced anaphylaxis by suppressing IgE
CC   production in response to the antigen challenge. The conjugates can
CC   also be used to combat pathogenic infection and to stimulate
CC   therapeutic angiogenesis to treat conditions in which localised blood
CC   flow plays a significant etiological role, e.g. retinopathies.
XX
SQ   Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

  Query Match             100.0%;  Score 22;  DB 19;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 0.034;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           ||||||||||||||||||||||
Db       1 tgactgtgaacgttcgagatga 22


RESULT   2
AAX36624
ID   AAX36624 standard; DNA; 22 BP.
XX
AC   AAX36624;
XX
DT   09-JUL-1999  (first entry)
XX
DE   ISS-ODN DY1018 nucleotide sequence.
XX
KW   Antigen-stimulated inflammation; immunostimulatory oligonucleotide;
KW   granulocyte-mediated tissue inflammation; Th2 type immune response;
KW   immune responsiveness modulation; idiopathic hypereosiniphilic syndrome;
KW   cutaneous basophil hypersensitivity; ISS-ODN; asthma; nasal polyposis;
KW   allergic rhinitis; atopic dermatitis; allergic conjunctivitis;
KW   eosinophilic fasciitis; therapy; ss.
XX
OS   Synthetic.
XX
PN   WO9911275-A2.
XX
PD   11-MAR-1999.
XX
PF   04-SEP-1998;   98WO-US18382.
XX
PR   05-SEP-1997;   97US-0927120.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Ray E;
XX
DR   WPI; 1999-312404/26.
XX
PT   Reducing antigen-stimulated granulocyte-mediated inflammation
XX
PS   Example 2; Page 30; 69pp; English.
XX
CC   This is the ISS-ODN DY1018 nucleotide sequence.
CC   The invention relates to a method for preventing or reducing
CC   antigen-stimulated, granulocyte-mediated tissue inflammation in a mammal,
CC   by administering an immunostimulatory oligonucleotide (ISS-ODN), where:
CC   (a) reduction in, or the absence of, a Th2 type immune response is
CC   measured; or (b) there is a reduction or absence of other clinical signs
CC   of inflammation in the host after antigen challenge. The method is used
CC   to reduce or suppress granulocyte-mediated inflammation in a host tissue,
CC   and to modulate the host's immune responsiveness to an antigen,
CC   particularly where the subject suffers from asthma, nasal polyposis,
CC   allergic rhinitis, atopic dermatitis, allergic conjunctivitis,
CC   eosinophilic fasciitis, idiopathic hypereosiniphilic syndrome, or
CC   cutaneous basophil hypersensitivity. Unlike prior art treatment by
CC   antigen immunisation, the method is an antigen-independent method,
CC   and avoids host production of both interleukin-4 (IL-4), which carries
CC   risk of anaphylaxis, and IL-5 which actually encourages granulocyte
CC   adhesion to endothelia.
XX
SQ   Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

  Query Match             100.0%;  Score 22;  DB 20;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 0.034;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           ||||||||||||||||||||||
Db       1 tgactgtgaacgttcgagatga 22


RESULT   3
AAV80097
ID   AAV80097 standard; DNA; 22 BP.
XX
AC   AAV80097;
XX
DT   12-MAR-1999  (first entry)
XX
DE   Immunomodulatory oligo comprising an ISS sequence.
XX
KW   Immunomodulatory; immunostimulatory; octanucleotide; immune regulation;
KW   ISS; cancer; allergy; asthma; hepatitis B infection; papillomavirus;
KW   human immunodeficiency virus; influenza; herpes; M. tuberculosis; ss;
KW   B. pertussis; malaria; plasmodia; Leishmania; Trypanosoma; Schistosoma.
XX
OS   Synthetic.
XX
PN   WO9855495-A2.
XX
PD   10-DEC-1998.
XX
PF   05-JUN-1998;   98WO-US11578.
XX
PR   06-JUN-1997;   97US-0048793.
XX
PA   (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI   Dina D,  Roman M,  Schwartz D;
XX
DR   WPI; 1999-059898/05.
XX
PT   Immunostimulatory oligonucleotides regulate the immune system - and
PT   contain an immune-stimulating octanucleotide sequence; for treating
PT   cancer, allergic and infectious diseases
XX
PS   Claim 5; Page 29; 63pp; English.
XX
CC   The invention relates to immunomodulatory oligonucleotides that comprise
CC   at least 1 immunostimulatory octanucleotide sequence (ISS) where the ISS
CC   sequences are selected from the group consisting of AACGTTCC, AACGTTCG,
CC   GACGTTCC, and GACGTTCG. The immunomodulatory sequences are used to treat
CC   patients needing immune regulation, such as those suffering from cancer,
CC   an allergic disease and asthma. They are also used to prevent infectious
```

```
CC   diseases such as influenza, herpes, hepatitis B, human immunodeficiency
CC   and papillomavirus, Hemophilus influenza, Mycobacterium tuberculosis and
CC   Bordetella pertussis, malarial plasmodia, Leishmania, Trypanosoma and
CC   Schistosoma. The immunomodulatory sequences are used to screen for human
CC   immunostimulatory activity by incubating macrophage cells and the
CC   oligonucleotide; and determining the relative amount of Th1-biased
CC   cytokines in the supernatant. Sequences AAV80096 to AAV80103 represent
CC   specific claimed examples of such immunomodulatory oligonucleotides.
XX
SQ   Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;


  Query Match             100.0%;  Score 22;  DB 20;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 0.034;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          ||||||||||||||||||||||
Db      1 tgactgtgaacgttcgagatga 22


RESULT   4
AAV80102
ID   AAV80102 standard; DNA; 22 BP.
XX
AC   AAV80102;
XX
DT   12-MAR-1999  (first entry)
XX
DE   Immunomodulatory oligo comprising an ISS sequence.
XX
KW   Immunomodulatory; immunostimulatory; octanucleotide; immune regulation;
KW   ISS: cancer; allergy; asthma; hepatitis B infection; papillomavirus;
KW   human immunodeficiency virus; influenza; herpes; M. tuberculosis; ss;
KW   B. pertussis; malaria; plasmodia; Leishmania; Trypanosoma; Schistosoma.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   modified_base    11
FT                    /*tag=  a
FT                    /note= "5-bromocytosine"
XX
PN   WO9855495-A2.
XX
PD   10-DEC-1998.
XX
PF   05-JUN-1998;   98WO-US11578.
XX
PR   06-JUN-1997;   97US-0048793.
XX
PA   (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI   Dina D,  Roman M,  Schwartz D;
XX
DR   WPI; 1999-059898/05.
XX
PT   Immunostimulatory oligonucleotides regulate the immune system - and
PT   contain an immune-stimulating octanucleotide sequence; for treating
PT   cancer, allergic and infectious diseases
XX
PS   Claim 23; Page 30; 63pp; English.
XX
CC   The invention relates to immunomodulatory oligonucleotides that comprise
CC   at least 1 immunostimulatory octanucleotide sequence (ISS) where the ISS
CC   sequences are selected from the group consisting of AACGTTCC, AACGTTCG,
CC   GACGTTCC, and GACGTTCG. The immunomodulatory sequences are used to treat
CC   patients needing immune regulation, such as those suffering from cancer,
CC   an allergic disease and asthma. They are also used to prevent infectious
CC   diseases such as influenza, herpes, hepatitis B, human immunodeficiency
CC   and papillomavirus, Hemophilus influenza, Mycobacterium tuberculosis and
CC   Bordetella pertussis, malarial plasmodia, Leishmania, Trypanosoma and
CC   Schistosoma. The immunomodulatory sequences are used to screen for human
CC   immunostimulatory activity by incubating macrophage cells and the
CC   oligonucleotide; and determining the relative amount of Th1-biased
CC   cytokines in the supernatant. Sequences AAV80096 to AAV80103 represent
CC   specific claimed examples of such immunomodulatory oligonucleotides.
XX
SQ   Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;


  Query Match             100.0%;  Score 22;  DB 20;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 0.034;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          ||||||||||||||||||||||
Db      1 tgactgtgaacgttcgagatga 22


RESULT   5
AAV80103
ID   AAV80103 standard; DNA; 22 BP.
XX
AC   AAV80103;
XX
DT   12-MAR-1999  (first entry)
XX
DE   Immunomodulatory oligo comprising an ISS sequence.
XX
KW   Immunomodulatory; immunostimulatory; octanucleotide; immune regulation;
KW   ISS: cancer; allergy; asthma; hepatitis B infection; papillomavirus;
KW   human immunodeficiency virus; influenza; herpes; M. tuberculosis; ss;
KW   B. pertussis; malaria; plasmodia; Leishmania; Trypanosoma; Schistosoma.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   modified_base    11
FT                    /*tag=  a
FT                    /note= "5-bromocytosine"
XX
PN   WO9855495-A2.
XX
PD   10-DEC-1998.
XX
PF   05-JUN-1998;   98WO-US11578.
XX
PR   06-JUN-1997;   97US-0048793.
XX
PA   (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI   Dina D,  Roman M,  Schwartz D;
XX
DR   WPI; 1999-059898/05.
XX
PT   Immunostimulatory oligonucleotides regulate the immune system - and
PT   contain an immune-stimulating octanucleotide sequence; for treating
PT   cancer, allergic and infectious diseases
XX
PS   Claim 24; Page 30; 63pp; English.
XX
CC   The invention relates to immunomodulatory oligonucleotides that comprise
CC   at least 1 immunostimulatory octanucleotide sequence (ISS) where the ISS
CC   sequences are selected from the group consisting of AACGTTCC, AACGTTCG,
CC   GACGTTCC, and GACGTTCG. The immunomodulatory sequences are used to treat
CC   patients needing immune regulation, such as those suffering from cancer,
CC   an allergic disease and asthma. They are also used to prevent infectious
CC   diseases such as influenza, herpes, hepatitis B, human immunodeficiency
CC   and papillomavirus, Hemophilus influenza, Mycobacterium tuberculosis and
CC   Bordetella pertussis, malarial plasmodia, Leishmania, Trypanosoma and
CC   Schistosoma. The immunomodulatory sequences are used to screen for human
CC   immunostimulatory activity by incubating macrophage cells and the
CC   oligonucleotide; and determining the relative amount of Th1-biased
CC   cytokines in the supernatant. Sequences AAV80096 to AAV80103 represent
CC   specific claimed examples of such immunomodulatory oligonucleotides.
```

```
XX
SQ   Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;


  Query Match             100.0%;  Score 22;  DB 20;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 0.034;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          ||||||||||||||||||||||
Db      1 tgactgtgaacgttcgagatga 22


RESULT   6
AAC64051
ID   AAC64051 standard; DNA; 22 BP.
XX
AC   AAC64051;
XX
DT   15-FEB-2001  (first entry)
XX
DE   Immunostimulatory CpG phosphorothioate oligodeoxynucleotide.
XX
KW   CpG oligodeoxynucleotide; phosphorothioate; immunostimulatory; ISS ODN;
KW   enhanced antigen presentation; antigen-presenting cell; APC;
KW   T-cell activation; tumour cell; tumour antigen; cancer immunotherapy;
KW   vaccine; ss.
XX
OS   Synthetic.
XX
PN   WO200062787-A1.
XX
PD   26-OCT-2000.
XX
PF   11-APR-2000; 2000WO-US09664.
XX
PR   15-APR-1999;   99US-0292278.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Raz E,  Martin-Orozco E;
XX
DR   WPI; 2000-679548/66.
XX
PT   Enhancing antigen-presentation capabilities of T-cells for cancer
PT   immunotherapy, by contacting cells with an immunostimulatory
PT   oligonucleotide  -
XX
PS   Example I; Page 18; 42pp; English.
XX
CC   The invention relates to a method of inducing activation of T-cells
CC   to respond to an antigen, comprising contacting antigen-presenting cells
CC   (APC) with an immunostimulatory oligodeoxynucleotide (ISS-ODN). The APCs
CC   thus treated have enhanced antigen presenting capabilities compared to
CC   antigen-activated APCs. APCs with enhanced antigen-presentation
CC   capabilities then present the antigen to T-cells. The method is useful
CC   for cancer immunotherapy. The ISS-ODN is used to enhance the tumour
CC   antigen presenting capacity of tumour cells, thereby inducing T-cell
CC   activation, and is therefore useful for treating tumours. Additionally,
CC   tumour cells treated with an ISS-ODN ex vivo are useful as vaccines.
CC   ISS-ODN treated APCs are induced to take up antigen through upregulation
CC   of Fc-receptor expression, to present antigen through upregulation of
CC   major histocompatibility complex (MHC) Class I and II expression and
CC   CD1d expression, to produce co-stimulatory factors (B7 and CD40), to
CC   provide cell-to-cell adhesion through upregulation of intercellular
CC   adhesion molecule (ICAM) expression, and to increase Th1 stimulatory
CC   cytokine production, all at levels greater than that achieved through
CC   contact of APC with antigen alone. The present sequence represents
CC   a phosphorothioate CpG ISS-ODN used in the exemplifications of the
CC   invention.
XX
SQ   Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;


  Query Match             100.0%;  Score 22;  DB 21;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 0.034;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          ||||||||||||||||||||||
Db      1 tgactgtgaacgttcgagatga 22


RESULT   7
AAA96253
ID   AAA96253 standard; DNA; 22 BP.
XX
AC   AAA96253;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Sequence of a stabilised oligonucleotide with antitumour activity.
XX
KW   Antitumour; immunostimulatory oligonucleotide; tumour; anaplasia;
KW   glioblastoma; medullablastoma; neuroblastoma; melanoma; carcinoma; ss.
XX
OS   Synthetic.
XX
PN   WO200056342-A2.
XX
PD   28-SEP-2000.
XX
PF   17-MAR-2000; 2000WO-FR00676.
XX
PR   19-MAR-1999;   99FR-0003433.
XX
PA   (ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.
PA   (INRM ) INST NAT SANTE & RECH MEDICALE.
XX
PI   Carpentier A;
XX
DR   WPI; 2000-602192/57.
XX
PT   Use of stabilized oligonucleotides as antitumor agents, particularly
PT   against nervous system tumors, have optimal activity and are not toxic
PT   -
XX
PS   Example 2; Page 16; 57pp; French.
XX
CC   The present sequence represents a stabilised oligonucleotide which has
CC   antitumour activity. The oligonucleotide comprises an octomer motif
CC   of the type 5'-purine-purine-CG-pyrimidine-pyrimidine-X-X-3', where
CC   the pair X-X is AT, AA, CT or TT. The oligonucleotides are
CC   immunostimulatory, and are not toxic. They may be adapted for use in
CC   animals or humans. The stabilised oligonucleotides are used for
CC   treating tumours, of any type and any degree of anaplasia, particularly
CC   human tumours in the peripheral or central nervous systems, specifically
CC   glioblastomas, medullablastomas, neuroblastomas, melanomas or carcinomas.
XX
SQ   Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;


  Query Match             100.0%;  Score 22;  DB 21;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 0.034;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          ||||||||||||||||||||||
Db      1 tgactgtgaacgttcgagatga 22


RESULT   8
AAA90458
ID   AAA90458 standard; DNA; 22 BP.
XX
AC   AAA90458;
```

XX
DT 10-JAN-2001 (first entry)
XX
DE CpG adjuvant oligonucleotide, SEQ ID NO:19.
XX
KW CpG oligonucleotide; CpG motif; adjuvant; microdroplet emulsion;
KW microemulsion; adsorbent microparticle; vaccine; Th1 immune response;
KW viral infection; bacterial infection; parasitic infection; HCV; HBV;
KW hepatitis C virus; hepatitis B virus; herpes simplex virus; HSV; HIV;
KW human immunodeficiency virus; cytomegalovirus; CMV; influenza virus;
KW rabies virus; cholera; diphtheria; tetanus; pertussis;
KW Helicobacter pylori; Haemophilus influenzae; malaria; ss.
XX
OS Synthetic.
XX
PN WO200050006-A2.
XX
PD 31-AUG-2000.
XX
PF 09-FEB-2000; 2000WO-US03331.
XX
PR 26-FEB-1999; 99US-0121858.
PR 29-JUL-1999; 99US-0146391.
PR 28-OCT-1999; 99US-0161997.
XX
PA (CHIR) CHIRON CORP.
XX
PI O'Hagan D, Ott GS, Donnelly J, Kazzaz J, Ugozzoli M, Singh M;
PI Barackman J;
XX
DR WPI; 2000-587123/55.
XX
PT Microemulsion having an adsorbent surface comprising a microdroplet
PT emulsion consisting of a metabolizable oil and an emulsifying agent
PT which is a detergent, useful as a vaccine to treat bacterial, viral,
PT and parasitic infection -
XX
PS Claim 17; Page 40; 95pp; English.
XX
CC The invention relates to a microdroplet emulsion (microemulsion) with an
CC adsorbent surface, and which comprises a metabolisable oil and an
CC emulsifying agent (a detergent). It also relates to a composition
CC comprising the microemulsion and a microparticle with an adsorbent
CC surface, where the microparticle comprises a polymer selected from a
CC poly(alpha-hydroxy acid), a polyhydroxy butyric acid, a
CC polycaprolactone, a polyorthoester, a polyanhydride, and a
CC polycyanoacrylate, and a second detergent. The surface of the
CC microparticles efficiently adsorb biologically active macromolecules such
CC as DNA, polypeptides, antigens, hormones, pharmaceuticals, enzymes,
CC mediators of transcription or translation, metabolic intermediates and
CC adjuvants. Additionally, a second biologically active molecule may be
CC encapsulated within the microparticle. The microemulsion can be used in
CC methods of immunising a host animal, particularly a human, against a
CC viral, bacterial or parasitic infection, and in methods of increasing a
CC Th1 immune response. The microemulsions (having the appropriate antigens
CC adsorbed) may be particularly used as vaccines for hepatitis C virus
CC (HCV), hepatitis B virus (HBV), herpes simplex virus (HSV), human
CC immunodeficiency virus (HIV), cytomegalovirus (CMV), influenza virus, and
CC rabies virus; the bacteria which cause cholera, diphtheria, tetanus and
CC pertussis; Helicobacter pylori and Haemophilus influenzae; and
CC malaria-causing parasites. Sequences AAA90447-A90467 represent Th1
CC lymphocyte stimulating oligonucleotides containing at least one CpG motif
CC which are claimed for use as adjuvants in the compositions of the
CC invention.
XX
SQ Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

Query Match 100.0%; Score 22; DB 21; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.034;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGACTGTGAACGTTCGAGATGA 22
|||||||||||||||||||||||
Db 1 tgactgtgaacgttcgagatga 22

RESULT 9
AAA14467
ID AAA14467 standard; DNA; 22 BP.
XX
AC AAA14467;
XX
DT 21-AUG-2000 (first entry)
XX
DE Immunostimulatory oligonucleotide (ISS-ODN) DY1018.
XX
KW Immunostimulatory oligonucleotide; adjuvant; mucosal immunity;
KW secretory immunoglobulin A production; sIgA; Th1 phenotype; ds.
XX
OS Synthetic.
XX
PN WO200020039-A1.
XX
PD 13-APR-2000.
XX
PF 15-SEP-1999; 99WO-US21203.
XX
PR 05-OCT-1998; 98US-0167039.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Raz E, Horner AA, Carson DA;
XX
DR WPI; 2000-303647/26.
XX
PT Immunostimulatory oligonucleotide adjuvant induces mucosal immunity to
PT an antigen in a mammalian host through production of secretory
PT immunoglobulin A -
XX
PS Claim 8; Page 21; 64pp; English.
XX
CC The invention relates to a method of inducing mucosal immunity to an
CC antigen in a mammalian host, including the the production of secretory
CC immunoglobulin A (sIgA). Immune protection in the mucosa (the principal
CC site of entry of most foreign antigens) is mediated by mucosa-associated
CC lymphoid tissue, epithelial and distinct B-cell, T-cell and accessory
CC cell sub-populations. The primary immune response which characterises
CC the induction of mucosal immunity to an antigen is sIgA production by
CC activated B-cells. The method comprises introducing an immunostimulatory
CC oligonucleotide (ISS-ODN) and the antigen into host mucosa, where the
CC ISS-ODN includes a core nucleotide sequence. The core nucleotide
CC sequence is 5'-Purine-Purine-C-G-Pyrimidine-Pyrimidine-3', specific
CC examples of which are AACGTT, AGCGTC and GACGTT (SEQ ID NOs 1-3). A
CC specific example of an ISS-ODN is DY1018 (AAA14467). The ISS-ODN is used
CC as an adjuvant with an antigen for stimulating mucosal immunity. The
CC level of sIgA production induced in the host is at least 3 times the
CC magnitude of sIgA production achievable in response to introduction of
CC antigen alone into the mucosal tissue and is equivalent or greater than
CC the magnitude of sIgA production achievable in response to introduction
CC of the antigen and cholera toxin adjuvant into the mucosal tissue. The
CC host immune response is stimulated to antigen specific IgA production,
CC biased towards the Th1 phenotype while antigen-induced IgE production is
CC avoided. The adjuvant has little or no known toxicity in mammals and its
CC efficacy is comparable to that of cholera toxin which is used as a
CC mucosal adjuvant. The present sequence represents the immunostimulatory
CC oligonucleotide DY1018.
XX
SQ Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

Query Match 100.0%; Score 22; DB 21; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.034;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGACTGTGAACGTTCGAGATGA 22

```
              ||||||||||||||||||||||
Db        1 tgactgtgaacgttcgagatga 22


RESULT 10
AAA38065
ID   AAA38065 standard; DNA; 22 BP.
XX
AC   AAA38065;
XX
DT   24-AUG-2000  (first entry)
XX
DE   Immunostimulatory sequence (ISS) #1.
XX
KW   Immunostimulatory sequence; ISS; immunomodulator; glycoprotein 120;
KW   gp120; human immunodeficiency virus; HIV; immune response; infection;
KW   development; ss.
XX
OS   Synthetic.
XX
PN   WO200021556-A1.
XX
PD   20-APR-2000.
XX
PF   08-OCT-1999;   99WO-US23677.
XX
PR   09-OCT-1998;   98US-0103733.
PR   07-OCT-1999;   99US-0415186.
XX
PA   (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI   Tighe H,  Raz E,  Schwartz D,  Takabayashi K;
XX
DR   WPI; 2000-317846/27.
XX
PT   Anti-HIV composition comprises immunostimulatory polynucleotides and
PT   HIV glycoprotein gp120 useful for modulating, stimulating an immune
PT   response against HIV in an HIV infected individual  -
XX
PS   Claim 3; Page 16; 65pp; English.
XX
CC   The present invention relates to an immunostimulatory composition
CC   comprising a human immunodeficiency virus (HIV) antigen, and an
CC   immunomodulatory polynucleotide comprising an immunostimulatory sequence
CC   (ISS). This sequence represents an ISS that can be used in the
CC   composition. An immunostimulatory composition which comprises a gp120
CC   conjugated to an immunomodulatory polynucleotide, or is proximately
CC   associated to it and not conjugated, is used for modulating or
CC   stimulating a specific immune response against gp120 in an individual by
CC   producing anti-gp120 antibodies or gp120 specific cytotoxic T cells. It
CC   is also used for suppressing or delaying development of HIV infection in
CC   an individual infected with HIV or an individual at risk of infection
CC   with HIV, respectively. It is also used for treating an individual
CC   infected with HIV in need of immune modulation.
XX
SQ   Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;


  Query Match             100.0%;  Score 22;  DB 21;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 0.034;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           ||||||||||||||||||||||
Db       1 tgactgtgaacgttcgagatga 22


RESULT 11
AAA38071
ID   AAA38071 standard; DNA; 22 BP.
XX
AC   AAA38071;
XX
DT   24-AUG-2000  (first entry)
XX
DE   Immunostimulatory sequence (ISS) #7.
XX
KW   Immunostimulatory sequence; ISS; immunomodulator; glycoprotein 120;
KW   gp120; human immunodeficiency virus; HIV; immune response; infection;
KW   development; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   modified_base   11
FT                   /*tag=  a
FT                   /mod_base= OTHER
FT                   /note= "5-Bromocytosine"
XX
PN   WO200021556-A1.
XX
PD   20-APR-2000.
XX
PF   08-OCT-1999;   99WO-US23677.
XX
PR   09-OCT-1998;   98US-0103733.
PR   07-OCT-1999;   99US-0415186.
XX
PA   (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI   Tighe H,  Raz E,  Schwartz D,  Takabayashi K;
XX
DR   WPI; 2000-317846/27.
XX
PT   Anti-HIV composition comprises immunostimulatory polynucleotides and
PT   HIV glycoprotein gp120 useful for modulating, stimulating an immune
PT   response against HIV in an HIV infected individual  -
XX
PS   Disclosure; Page 17; 65pp; English.
XX
CC   The present invention relates to an immunostimulatory composition
CC   comprising a human immunodeficiency virus (HIV) antigen, and an
CC   immunomodulatory polynucleotide comprising an immunostimulatory sequence
CC   (ISS). This sequence represents an ISS that can be used in the
CC   composition. An immunostimulatory composition which comprises a gp120
CC   conjugated to an immunomodulatory polynucleotide, or is proximately
CC   associated to it and not conjugated, is used for modulating or
CC   stimulating a specific immune response against gp120 in an individual by
CC   producing anti-gp120 antibodies or gp120 specific cytotoxic T cells. It
CC   is also used for suppressing or delaying development of HIV infection in
CC   an individual infected with HIV or an individual at risk of infection
CC   with HIV, respectively. It is also used for treating an individual
CC   infected with HIV in need of immune modulation.
XX
SQ   Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;


  Query Match             100.0%;  Score 22;  DB 21;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 0.034;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           ||||||||||||||||||||||
Db       1 tgactgtgaacgttcgagatga 22


RESULT 12
AAA38072
ID   AAA38072 standard; DNA; 22 BP.
XX
AC   AAA38072;
XX
DT   24-AUG-2000  (first entry)
XX
DE   Immunostimulatory sequence (ISS) #7.
XX
```

```
KW   Immunostimulatory sequence; ISS; immunomodulator; glycoprotein 120;
KW   gp120; human immunodeficiency virus; HIV; immune response; infection;
KW   development; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   modified_base   11
FT                   /*tag=  a
FT                   /mod_base= OTHER
FT                   /note= "5-Bromocytosine"
FT   modified_base   15
FT                   /*tag=  b
FT                   /mod_base= OTHER
FT                   /note= "5-Bromocytosine"
XX
PN   WO200021556-A1.
XX
PD   20-APR-2000.
XX
PF   08-OCT-1999;   99WO-US23677.
XX
PR   09-OCT-1998;   98US-0103733.
PR   07-OCT-1999;   99US-0415186.
XX
PA   (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI   Tighe H,  Raz E,  Schwartz D,  Takabayashi K;
XX
DR   WPI; 2000-317846/27.
XX
PT   Anti-HIV composition comprises immunostimulatory polynucleotides and
PT   HIV glycoprotein gp120 useful for modulating, stimulating an immune
PT   response against HIV in an HIV infected individual
XX
PS   Disclosure; Page 17; 65pp; English.
XX
CC   The present invention relates to an immunostimulatory composition
CC   comprising a human immunodeficiency virus (HIV) antigen, and an
CC   immunomodulatory polynucleotide comprising an immunostimulatory sequence
CC   (ISS). This sequence represents an ISS that can be used in the
CC   composition. An immunostimulatory composition which comprises a gp120
CC   conjugated to an immunomodulatory polynucleotide, or is proximately
CC   associated to it and not conjugated, is used for modulating or
CC   stimulating a specific immune response against gp120 in an individual by
CC   producing anti-gp120 antibodies or gp120 specific cytotoxic T cells. It
CC   is also used for suppressing or delaying development of HIV infection in
CC   an individual infected with HIV or an individual at risk of infection
CC   with HIV, respectively. It is also used for treating an individual
CC   infected with HIV in need of immune modulation.
XX
SQ   Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

  Query Match           100.0%;  Score 22;  DB 21;  Length 22;
  Best Local Similarity  100.0%;  Pred. No. 0.034;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           ||||||||||||||||||||||
Db       1 tgactgtgaacgttcgagatga 22


RESULT  13
AAZ55876
ID   AAZ55876 standard; DNA; 22 BP.
XX
AC   AAZ55876;
XX
DT   10-APR-2000  (first entry)
XX
DE   Immunomodulatory oligonucleotide SEQ ID NO: 1.
XX
KW   Immunomodulation; immunostimulatory sequence; adjuvant;
KW   Th1 immune response; cytotoxic T-cell; cytokine; cancer; allergy;
KW   asthma; immunocontraception; ss.
XX
OS   Mus musculus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   modified_base   1..22
FT                   /*tag=  a
FT                   /note= "Phosphorothioate linkages"
FT   misc_feature    9..16
FT                   /*tag=  b
FT                   /note= "Immunostimulatory sequence (ISS)"
XX
PN   WO9962923-A2.
XX
PD   09-DEC-1999.
XX
PF   04-JUN-1999;   99WO-US12538.
XX
PR   05-JUN-1998;   98US-0088310.
PR   01-JUN-1999;   99US-0324191.
XX
PA   (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI   Schwartz D;
XX
DR   WPI; 2000-105687/09.
XX
PT   Novel immunomodulatory oligonucleotide used to induce a Th1-type immune
PT   response, e.g. to tumor antigens
XX
PS   Example 1; Page 35; 54pp; English.
XX
CC   Sequences AAZ55876-Z55877 and AAZ55880-Z55886 represent immunomodulatory
CC   oligonucleotides comprising an immunostimulatory sequence (ISS, e.g.,
CC   AACGTC, AACGTT, AGCGTC, AGCGCT, AGCGTT, GACGTC, GACGTT, GGCGTT,
CC   AACGTTCC and GACGTTCC). The invention relates to oligonucleotides
CC   comprising one or more ISSs, where the ISS comprises at least
CC   one modified cytosine with an electron-withdrawing moiety at
CC   position C-5 or C-6 of the base. Sequences AAZ55877 and AAZ55880-Z55886
CC   contain ISSs comprising at least one bromocytosine, whereas sequence
CC   AAZ55876 contains an unmodified ISS. The immunomodulatory
CC   oligonucleotides have an adjuvant-like effect; when formulated with an
CC   antigen, the oligonucleotides stimulate production of Th1-type cytokines,
CC   and induce a Th1-type immune response (activation of cytotoxic T cells),
CC   while simultaneously downregulating the Th2-type response. The Th1
CC   response is particularly effective for control of viruses and
CC   intracellular parasites. The immunomodulatory oligonucleotides are used,
CC   particularly when formulated with an antigen or a facilitator, for
CC   modulating immune responses. Such compositions may be used in tumour
CC   therapy, in treatment of allergy (including asthma), for inducing a
CC   vigorous cellular response (against a virus, bacterium, fungus or
CC   protozoan), and also in contraceptive vaccines based on sperm antigens.
XX
SQ   Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

  Query Match           100.0%;  Score 22;  DB 21;  Length 22;
  Best Local Similarity  100.0%;  Pred. No. 0.034;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           ||||||||||||||||||||||
Db       1 tgactgtgaacgttcgagatga 22


RESULT  14
AAH42533
ID   AAH42533 standard; DNA; 22 BP.
XX
AC   AAH42533;
```

```
XX
DT   01-OCT-2001  (first entry)
XX
DE   Phosphorothioate beta-gal/immunostimulatory oligonucleotide.
XX
KW   Anaphylactic hypersensitivity; immunomodulatory nucleic acid; vaccine;
KW   anaphylaxis-associated symptom; IgE; histamine; phosphorothioate; ss.
XX
OS   Synthetic.
XX
PN   WO200145750-A1.
XX
PD   28-JUN-2001.
XX
PF   20-DEC-2000; 2000WO-US35064.
XX
PR   21-DEC-1999;   99US-0171830.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Raz E,  Horner AA;
XX
DR   WPI; 2001-475812/51.
XX
PT   Reducing risk of anaphylactic hypersensitivity response to an allergen
PT   in a subject, by administering an immunomodulating nucleic acid
PT   molecule comprising a specific sequence  -
XX
PS   Example 1; Page 22; 39pp; English.
XX
CC   The specification describes a method for reducing a symptom associated
CC   with anaphylactic hypersensitivity or risk of anaphylactic response in
CC   a subject. The method comprises administering to an individual a
CC   nucleic acid molecule comprising an immunomodulatory nucleic acid
CC   molecule (INA) comprising the sequence 5'-C-G-3' to reduce
CC   anaphylaxis-associated symptom. The method is useful for reducing a
CC   symptom associated with anaphylactic hypersensitivity, including
CC   elevated IgE level, elevated histamine level, constriction of the
CC   airways and difficult breathing which can lead to anaphylactic reaction
CC   or anaphylactic shock, thereby reducing the risk of death. The present
CC   sequence represents a beta-gal/immunostimulatory sequence, which was
CC   used as a vaccine to protect against the development of anaphylactic
CC   hypersensitivity.
XX
SQ   Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;


  Query Match             100.0%;  Score 22;  DB 22;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 0.034;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           ||||||||||||||||||||||
Db       1 tgactgtgaacgttcgagatga 22


RESULT  15
AAH73439
ID   AAH73439 standard; DNA; 22 BP.
XX
AC   AAH73439;
XX
DT   01-OCT-2001  (first entry)
XX
DE   Immunomodulatory nucleic acid.
XX
KW   G3PDH gene; immunomodulatory oligonucleotide; infection; mycobacterium;
KW   intracellular pathogen; anti-pathogenic; ss.
XX
OS   Unidentified.
XX
PN   WO200155341-A2.
XX
PD   02-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US03029.
XX
PR   31-JAN-2000; 2000US-0179353.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Raz E,  Kornbluth R,  Catanzaro A,  Hayashi T,  Carson DA;
XX
DR   WPI; 2001-483234/52.
XX
PT   Treating infection of intracellular pathogen e.g., Mycobacterium, in a
PT   subject, involves administering immunomodulatory nucleic acid molecule
PT   to inhibit intracellular replication of intracellular pathogen  -
XX
PS   Examples; Page 26; 54pp; English.
XX
CC   The present invention describes a method of treating an infection caused
CC   by an intracellular pathogen, involving administering to the patient an
CC   immunomodulatory nucleic acid and an anti-pathogenic agent. This is
CC   particularly useful in the treatment of mycobacterial infections. The
CC   present sequence is an immunomodulatory nucleic acid described in the
CC   exemplification of the invention.
XX
SQ   Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;


  Query Match             100.0%;  Score 22;  DB 22;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 0.034;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           ||||||||||||||||||||||
Db       1 tgactgtgaacgttcgagatga 22


RESULT  16
AAH44109
ID   AAH44109 standard; DNA; 22 BP.
XX
AC   AAH44109;
XX
DT   12-SEP-2001  (first entry)
XX
DE   5' terminal NH2 group and a 3' terminal rhodamine moiety oligonucleotide.
XX
KW   Peptide nucleic acid; intracellular protein delivery; cationic lipid;
KW   PNA; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   modified_base    1
FT                    /*tag=  a
FT                    /mod_base= OTHER
FT                    /note= "T has been modified at the 5' terminal with
FT                          an NH2 group"
FT   modified_base    22
FT                    /*tag=  b
FT                    /mod_base= OTHER
FT                    /note= "A has been modified at the 3' terminal with
FT                          rhodamine"
XX
PN   WO200143778-A1.
XX
PD   21-JUN-2001.
XX
PF   15-DEC-2000; 2000WO-US33969.
XX
PR   17-DEC-1999;   99US-0172441.
XX
PA   (GENE-) GENE THERAPY SYSTEMS INC.
```

```
XX
PI   Felgner PL,  Zelphati O;
XX
DR   WPI; 2001-398080/42.
XX
PT   Composition useful for intracellular delivery of a protein, comprises a
PT   protein in operative association with a cationic intracellular delivery
PT   vehicle comprising a cationic lipid, which is adapted to fuse with a
PT   cell membrane -
XX
PS   Example 3; Page 18; 33pp; English.
XX
CC   The present invention describes a composition (I) for intracellular
CC   delivery of a protein, comprising a protein in operative association
CC   with a cationic intracellular delivery vehicle comprising a cationic
CC   lipid, where the intracellular delivery vehicle is adapted to fuse with
CC   a cell membrane, therefore effecting intracellular delivery of the
CC   associated protein. also described is a method for delivering a protein
CC   to a cell involving providing the protein associated with a cationic
CC   lipid in such a manner so as to form an intracellular delivery
CC   composition, and contacting the delivery composition with a cell
CC   membrane of a cell, such that the cationic lipid forms an association
CC   with a cell membrane and delivers the protein into the cell. (I) is
CC   useful in the preparation of a medicament for intracellular delivery of
CC   a therapeutic or prophylactic protein. (I) is useful for delivering
CC   antibodies to intracellular proteins to neutralise their activity, and
CC   to introduce therapeutically useful, proteins, peptides or small
CC   molecules. (I) is useful for the in vitro or in vivo delivery of
CC   antibodies or peptides which block the function of specific intracellular
CC   proteins and affect cellular metabolism, cell viability or virus
CC   replication. (I) is useful for delivering any protein of interest,
CC   including therapeutically useful proteins (e.g. tumour suppressor
CC   proteins, cystic fibrosis transmembrane regulator (CFTR), adenosine
CC   deaminase (ADA), hexoseaminidase A, peptides, wild type protein
CC   counterparts of mutant proteins and cell surface receptors) such as
CC   those for cytokines (e.g., interleukins, interferons, colony stimulating
CC   factors) and peptide hormones. The present sequence represents a peptide
CC   nucleic acid (PNA) oligonucleotide which is used in an example from the
CC   present invention for intracellular delivery of proteins.
XX
SQ   Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

  Query Match             100.0%;  Score 22;  DB 22;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 0.034;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           ||||||||||||||||||||||
Db       1 tgactgtgaacgttcgagatga 22


RESULT  17
AAH41573
ID   AAH41573 standard; DNA; 22 BP.
XX
AC   AAH41573;
XX
DT   24-AUG-2001  (first entry)
XX
DE   Immunostimulatory sequence (ISS) SEQ ID NO:1.
XX
KW   Immunostimulatory sequence; ISS; immunomodulatory; immune response;
KW   antigen; antiallergic; modulation; Th1 lymphocyte stimulation; allergy;
KW   Th1-associated cytokine; Th2 lymphocyte suppression; cytokine; ss.
XX
OS   Synthetic.
XX
PN   WO200135991-A2.
XX
PD   25-MAY-2001.
XX
PF   15-NOV-2000; 2000WO-US31385.
XX
PR   15-NOV-1999;   99US-0165467.
PR   14-NOV-2000; 2000US-0713136.
XX
PA   (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI   Tuck S,  Van Nest G;
XX
DR   WPI; 2001-329209/34.
XX
PT   Populations of conjugate molecules comprising polynucleotide
PT   immunostimulatory sequences polynucleotides and antigens, useful for
PT   controlling immune responses -
XX
PS   Example 1; Page 30; 97pp; English.
XX
CC   The present invention describes immunomodulatory populations ((I) and
CC   (II)) of conjugate molecules (CMs) comprising immunostimulatory sequences
CC   (ISS) of polynucleotides and antigens. The extent of conjugation affects
CC   the immunological properties (e.g. the extent of antigen-specific
CC   antibody formation, including Th1-associated antibody formation) so the
CC   conjugates are used for altering the type and extent of immune response.
CC   (I) and (II) have immunomodulatory, immunosuppressive and antiallergic
CC   activities, and can be used in the modulation of immune responses via
CC   the stimulation of Th1 lymphocytes and Th1-associated cytokines, and
CC   suppression of Th2 lymphocytes and cytokines. The populations ((I) and
CC   (II)) of conjugate molecules may be used for modulating immune responses
CC   in individuals e.g. for the treatment of an allergic condition. (I) and
CC   (II) may be used to modulate immune responses and therefore prevent
CC   potentially harmful reactions to antigens. The present sequence
CC   represents an ISS polynucleotide which is used in the exemplification
CC   of the present invention.
XX
SQ   Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

  Query Match             100.0%;  Score 22;  DB 22;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 0.034;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           ||||||||||||||||||||||
Db       1 tgactgtgaacgttcgagatga 22


RESULT  18
AAH20403
ID   AAH20403 standard; DNA; 22 BP.
XX
AC   AAH20403;
XX
DT   03-AUG-2001  (first entry)
XX
DE   CpG motif containing oligonucleotide SEQ ID #21.
XX
KW   Immune system stimulator; CpG motif; CpG receptor; CpG-R; antibacterial;
KW   immune response; vaccine adjuvant; tumour immunotherapy; allergy;
KW   anti-inflammatory; cystic fibrosis; sepsis; heart disease; chlamydia;
KW   inflammatory bowel disease; arthritis; multiple sclerosis; ss.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   modified_base   1..22
FT                   /*tag=  a
FT                   /mod_base= OTHER
FT                   /note= "Phosphorothioate internucleoside linkages"
XX
PN   WO200132877-A2.
XX
PD   10-MAY-2001.
XX
PF   01-NOV-2000; 2000WO-US41735.
```

```
XX
PR   02-NOV-1999;   99US-0163157.
PR   24-NOV-1999;   99US-0167389.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Mackichan ML;
XX
DR   WPI; 2001-343486/36.
XX
PT   Novel CpG receptor and nucleic acid molecule encoding the receptor, for
PT   modulating immune response and for identifying compounds of therapeutic
PT   use which bind and/or modulate the activity of the receptor -
XX
PS   Example 1; Page 14; 41pp; English.
XX
CC   Unmethylated CG dinucleotide sequences are commonly found in bacterial
CC   DNA, and have been found to stimulate the innate immune system. Natural
CC   killer and T cells are activated by exposure to oligonucleotides
CC   containing CpG motifs. Oligonucleotides containing CpG motifs can be used
CC   as adjuvants in vaccines. The present invention relates to a CpG
CC   receptor. The CpG receptor contains a Toll homology domain (THD). The
CC   Toll receptor family are associated with responses to pathogens. CpG
CC   oligonucleotides may act as stimulators of various immune responses. The
CC   CpG receptor or cells expressing the receptor are useful for identifying
CC   a compound which binds to or modulates an activity of the CpG receptor.
CC   The compounds are useful in e.g. vaccine adjuvants promoting
CC   cell-mediated immune responses, antibacterials, (e.g. protection from
CC   Listeria infection), tumour immunotherapy, allergy treatment, (e.g.
CC   suppressing IgE in human PBMC, shifting from Th2 to Th1) and as
CC   anti-inflammatory agents (e.g. for use in cystic fibrosis, sepsis, heart
CC   disease, chlamydia, inflammatory bowel disease, arthritis and multiple
CC   sclerosis). The present sequence represents a CpG motif containing
CC   oligonucleotide used in examples demonstrating that CpG oligonucleotides
CC   can activate the MAPK pathways and NF-kappaB.
XX
SQ   Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

  Query Match             100.0%;  Score 22;  DB 22;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 0.034;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           ||||||||||||||||||||||
Db       1 tgactgtgaacgttcgagatga 22


RESULT  19
AAF77040
ID   AAF77040 standard; DNA; 22 BP.
XX
AC   AAF77040;
XX
DT   15-MAY-2001  (first entry)
XX
DE   Immunomodulatory DNA.
XX
KW   Modulate; immune; antigen; immunostimulatory; ds.
XX
OS   Synthetic.
XX
PN   WO200112223-A2.
XX
PD   22-FEB-2001.
XX
PF   18-AUG-2000; 2000WO-US22835.
XX
PR   19-AUG-1999;   99US-0149768.
XX
PA   (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI   Van Nest G;
XX
DR   WPI; 2001-211136/21.
XX
PT   Modulating immune response to a second antigen in humans involves
PT   administering an immunostimulatory polynucleotide comprising an
PT   immunostimulatory sequence and a first antigen -
XX
PS   Claim 31; Page 15; 63pp; English.
XX
CC   The present invention relates to modulating an immune response to
CC   a second antigen in an individual, involving
CC   administering to the individual an immunomodulatory polynucleotide
CC   comprising an immunostimulatory sequence (ISS) and a first antigen.
XX
SQ   Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

  Query Match             100.0%;  Score 22;  DB 22;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 0.034;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           ||||||||||||||||||||||
Db       1 tgactgtgaacgttcgagatga 22


RESULT  20
AAF29800
ID   AAF29800 standard; DNA; 22 BP.
XX
AC   AAF29800;
XX
DT   12-APR-2001  (first entry)
XX
DE   Cholera toxin immunostimulatory nucleotide sequence.
XX
KW   Immunostimulatory nucleotide sequence; immune response; cancer;
KW   antibody production; IFNgamma release; CTL activity; Th1 response;
KW   infection; allergy; ds.
XX
OS   Unidentified.
XX
PN   WO200102007-A1.
XX
PD   11-JAN-2001.
XX
PF   30-JUN-2000; 2000WO-US18229.
XX
PR   02-JUL-1999;   99US-0347343.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Raz E,  Kobayashi H;
XX
DR   WPI; 2001-138066/14.
XX
PT   Enhancing immune response against pathogen or antigen associated with
PT   infectious diseases, an allergen or cancer, involves administering
PT   immunostimulatory nucleotide sequence prior to antigen exposure -
XX
PS   Example 1; Page 14; 47pp; English.
XX
CC   The present invention describes a method for enhancing an immune response
CC   to a substance, comprising administering an immunostimulatory nucleotide
CC   sequence to a subject prior to exposure to the substance. This can be
CC   used to enhance antibody production, IFNgamma release, CTL activity and
CC   Th1 related effects. The method can be used in the prevention and
CC   treatment of allergies, cancer and infections.
XX
SQ   Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

  Query Match             100.0%;  Score 22;  DB 22;  Length 22;
```

```
  Best Local Similarity   100.0%;  Pred. No. 0.034;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           ||||||||||||||||||||||
Db       1 tgactgtgaacgttcgagatga 22


RESULT  21
AAC82107
ID   AAC82107 standard; DNA; 22 BP.
XX
AC   AAC82107;
XX
DT   07-MAR-2001  (first entry)
XX
DE   Oligonucleotide ODNOct DNA SEQ ID NO 2.
XX
KW   Immunogenic; human immunodeficiency virus; immunostimulatory sequence;
KW   ISS; beta-chemokine; anti-HIV; AIDS; Th1 immune response; primer;
KW   HIV-specific cytotoxic T lymphocyte response; phosphorothioate; ss.
XX
OS   Synthetic.
XX
PN   WO200067787-A2.
XX
PD   16-NOV-2000.
XX
PF   05-MAY-2000; 2000WO-US12495.
XX
PR   06-MAY-1999;   99US-0132762.
PR   25-AUG-1999;   99US-0150667.
XX
PA   (IMMU-) IMMUNE RESPONSE CORP.
XX
PI   Moss RB;
XX
DR   WPI; 2001-031804/04.
XX
PT   Human immunodeficiency virus (HIV) compositions useful for immunizing
PT   and inhibiting AIDS in mammals, comprises HIV devoid of outer envelope
PT   protein and an immunostimulatory nucleic acid sequence  -
XX
PS   Example I; Page 26; 64pp; English.
XX
CC   This invention describes a novel immunogenic composition (I), comprising
CC   a whole-killed human immunodeficiency virus (HIV) devoid of outer
CC   envelope protein gp120, an isolated nucleic acid molecule containing an
CC   immunostimulatory sequence (ISS) and an adjuvant, which enhances
CC   beta-chemokine levels in a mammal. The products of the invention have
CC   anti-HIV activity. (I) is useful for immunizing and for inhibiting AIDS
CC   in a mammal. The mammal can be a primate such as a human, (HIV
CC   seronegative or seropositive humans) or a rodent, in particular the
CC   primate is a pregnant mother or an infant. (I) can induce potent Th1
CC   immune responses against a broad spectrum of HIV epitopes and provides a
CC   strong HIV-specific cytotoxic T lymphocyte response.
XX
SQ   Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;


  Query Match             100.0%;  Score 22;  DB 22;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 0.034;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           ||||||||||||||||||||||
Db       1 tgactgtgaacgttcgagatga 22


RESULT  22
AAA92377
ID   AAA92377 standard; DNA; 22 BP.
XX
AC   AAA92377;
XX
DT   12-JAN-2001  (first entry)
XX
DE   CG motif and CFA containing oligonucleotide SEQ ID NO:19.
XX
KW   CG motif; complete Freund's adjuvant; phosphorothioate; immunogenic;
KW   Neisseria antigen; Neisseria meningitidis; Neisseria gonorrhoeae;
KW   bactericidal; antibacterial; vaccine; immunostimulatory; infection;
KW   immune response; ss.
XX
OS   Neisseria sp.
XX
FH   Key             Location/Qualifiers
FT   modified_base   1..22
FT                   /*tag=  a
FT                   /note= "preferably contains at least one
FT                          phosphorothioate bond"
XX
PN   WO200050075-A2.
XX
PD   31-AUG-2000.
XX
PF   09-FEB-2000; 2000WO-IB00176.
XX
PR   26-FEB-1999;   99US-0121792.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Grandi G,  Rappuoli R,  Giuliani MM,  Pizza M;
XX
DR   WPI; 2001-015529/02.
XX
PT   Immunogenic composition useful for stimulating an immune response in a
PT   mammal against Neisseria infection, comprises Neisseria antigen and an
PT   adjuvant composition comprising an oligonucleotide with a CG motif -
XX
PS   Claim 19; Page 9; 39pp; English.
XX
CC   The present invention describes an immunogenic composition (I)
CC   comprising a Neisseria antigen and an adjuvant composition comprising an
CC   oligonucleotide comprising at least 1 CG motif. Also described is an
CC   adjuvant composition (II) comprising an oligonucleotide which comprises
CC   at least 1 CG motif and a complete Freund's adjuvant (CFA), where the
CC   oligonucleotide preferably comprises at least one phosphorothioate bond.
CC   AAA92359 to AAA92385 represent specifically claimed oligonucleotides of
CC   the present invention. (I) is useful for stimulating an immune response
CC   in a mammal, preferably a human, against Neisseria infection, preferably
CC   Neisseria meningitidis infection and in the manufacture of a medicament
CC   for inducing a protective immune response in a mammal.
XX
SQ   Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;


  Query Match             100.0%;  Score 22;  DB 22;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 0.034;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           ||||||||||||||||||||||
Db       1 tgactgtgaacgttcgagatga 22


RESULT  23
AAF77046
ID   AAF77046 standard; DNA; 22 BP.
XX
AC   AAF77046;
XX
DT   15-MAY-2001  (first entry)
XX
DE   Immunostimulatory DNA #6.
XX
```

```
KW   Modulate; immune; antigen; immunostimulatory; ds.
XX
OS   Synthetic.
XX
PN   WO200112223-A2.
XX
PD   22-FEB-2001.
XX
PF   18-AUG-2000; 2000WO-US22835.
XX
PR   19-AUG-1999;   99US-0149768.
XX
PA   (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI   Van Nest G;
XX
DR   WPI; 2001-211136/21.
XX
PT   Modulating immune response to a second antigen in humans involves
PT   administering an immunostimulatory polynucleotide comprising an
PT   immunostimulatory sequence and a first antigen -
XX
PS   Disclosure; Page 15; 63pp; English.
XX
CC   The present invention relates to modulating an immune response to
CC   a second antigen in an individual, involving
CC   administering to the individual an immunomodulatory polynucleotide
CC   comprising an immunostimulatory sequence (ISS) and a first antigen.
XX
SQ   Sequence 22 BP; 6 A; 2 C; 7 G; 6 T; 1 other;
```

```
  Query Match             96.4%;  Score 21.2;  DB 22;  Length 22;
  Best Local Similarity   95.5%;  Pred. No. 0.087;
  Matches   21;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           ||||||||||:||||||||||
Db       1 tgactgtgaabgttcgagatga 22
```

```
RESULT  24
AAZ55880
ID   AAZ55880 standard; DNA; 22 BP.
XX
AC   AAZ55880;
XX
DT   10-APR-2000  (first entry)
XX
DE   Immunomodulatory oligonucleotide SEQ ID NO: 5.
XX
KW   Immunomodulation; immunostimulatory sequence; adjuvant;
KW   Th1 immune response; cytotoxic T-cell; cytokine; cancer; allergy;
KW   asthma; immunocontraception; 5-bromocytosine; ss.
XX
OS   Mus musculus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   modified_base    1..22
FT                    /*tag= a
FT                    /note= "Phosphorothioate linkages"
FT   misc_feature     9..16
FT                    /*tag= b
FT                    /note= "Immunostimulatory sequence (ISS)"
FT   modified_base    11
FT                    /*tag= c
FT                    /mod_base= OTHER
FT                    /note= "5-bromocytosine"
XX
PN   WO9962923-A2.
XX
PD   09-DEC-1999.
XX
PF   04-JUN-1999;   99WO-US12538.
XX
PR   05-JUN-1998;   98US-0088310.
PR   01-JUN-1999;   99US-0324191.
XX
PA   (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI   Schwartz D;
XX
DR   WPI; 2000-105687/09.
XX
PT   Novel immunomodulatory oligonucleotide used to induce a Th1-type immune
PT   response, e.g. to tumor antigens -
XX
PS   Claim 30; Page 35; 54pp; English.
XX
CC   Sequences AAZ55876-Z55877 and AAZ55880-Z55886 represent immunomodulatory
CC   oligonucleotides comprising an immunostimulatory sequence (ISS, e.g.,
CC   AACGTC, AACGTT, AGCGTC, AGCGCT, AGCGTT, GACGTC, GACGTT, GGCGTT,
CC   AACGTTCC and GACGTTCC). The invention relates to oligonucleotides
CC   comprising one or more ISSs, where the ISS comprises at least
CC   one modified cytosine with an electron-withdrawing moiety at
CC   position C-5 or C-6 of the base. Sequences AAZ55877 and AAZ55880-Z55886
CC   contain ISSs comprising at least one bromocytosine, whereas sequence
CC   AAZ55876 contains an unmodified ISS. The immunomodulatory
CC   oligonucleotides have an adjuvant-like effect; when formulated with an
CC   antigen, the oligonucleotides stimulate production of Th1-type cytokines,
CC   and induce a Th1-type immune response (activation of cytotoxic T cells),
CC   while simultaneously downregulating the Th2-type response. The Th1
CC   response is particularly effective for control of viruses and
CC   intracellular parasites. The immunomodulatory oligonucleotides are used,
CC   particularly when formulated with an antigen or a facilitator, for
CC   modulating immune responses. Such compositions may be used in tumour
CC   therapy, in treatment of allergy (including asthma), for inducing a
CC   vigorous cellular response (against a virus, bacterium, fungus or
CC   protozoan), and also in contraceptive vaccines based on sperm antigens.
XX
SQ   Sequence 22 BP; 6 A; 2 C; 7 G; 6 T; 1 other;
```

```
  Query Match             95.5%;  Score 21;  DB 21;  Length 22;
  Best Local Similarity   95.5%;  Pred. No. 0.11;
  Matches   21;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           |||||||||| ||||||||||
Db       1 tgactgtgaangttcgagatga 22
```

```
RESULT  25
AAH41579
ID   AAH41579 standard; DNA; 22 BP.
XX
AC   AAH41579;
XX
DT   24-AUG-2001  (first entry)
XX
DE   Immunostimulatory sequence (ISS) SEQ ID NO:7.
XX
KW   Immunostimulatory sequence; ISS; immunomodulatory; immune response;
KW   antigen; antiallergic; modulation; Th1 lymphocyte stimulation; allergy;
KW   Th1-associated cytokine; Th2 lymphocyte suppression; cytokine; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   modified_base    11
FT                    /*tag= a
FT                    /mod_base= "OTHER"
FT                    /note= "5-bromocytosine"
XX
PN   WO200135991-A2.
```

XX
PD 25-MAY-2001.
XX
PF 15-NOV-2000; 2000WO-US31385.
XX
PR 15-NOV-1999; 99US-0165467.
PR 14-NOV-2000; 2000US-0713136.
XX
PA (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI Tuck S, Van Nest G;
XX
DR WPI; 2001-329209/34.
XX
PT Populations of conjugate molecules comprising polynucleotide
PT immunostimulatory sequences polynucleotides and antigens, useful for
PT controlling immune responses -
XX
PS Disclosure; Page 30; 97pp; English.
XX
CC The present invention describes immunomodulatory populations ((I) and
CC (II)) of conjugate molecules (CMs) comprising immunostimulatory sequences
CC (ISS) of polynucleotides and antigens. The extent of conjugation affects
CC the immunological properties (e.g. the extent of antigen-specific
CC antibody formation, including Th1-associated antibody formation) so the
CC conjugates are used for altering the type and extent of immune response.
CC (I) and (II) have immunomodulatory, immunosuppressive and antiallergic
CC activities, and can be used in the modulation of immune responses via
CC the stimulation of Th1 lymphocytes and Th1-associated cytokines, and
CC suppression of Th2 lymphocytes and cytokines. The populations ((I) and
CC (II)) of conjugate molecules may be used for modulating immune responses
CC in individuals e.g. for the treatment of an allergic condition. (I) and
CC (II) may be used to modulate immune responses and therefore prevent
CC potentially harmful reactions to antigens. The present sequence
CC represents an ISS polynucleotide which is used in the exemplification
CC of the present invention.
XX
SQ Sequence 22 BP; 6 A; 2 C; 7 G; 6 T; 1 other;

Query Match 95.5%; Score 21; DB 22; Length 22;
Best Local Similarity 95.5%; Pred. No. 0.11;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 TGACTGTGAACGTTCGAGATGA 22
           ||||||||| ||||||||||
Db       1 tgactgtgaangttcgagatga 22
```

RESULT 26
AAV80105/c
ID AAV80105 standard; DNA; 22 BP.
XX
AC AAV80105;
XX
DT 12-MAR-1999 (first entry)
XX
DE Oligo used in experiments for stimulation of cytokine production.
XX
KW Immunomodulatory; immunostimulatory; octanucleotide; immune regulation;
KW ISS: cancer; allergy; asthma; hepatitis B infection; papillomavirus;
KW human immunodeficiency virus; influenza; herpes; M. tuberculosis; ss;
KW B. pertussis; malaria; plasmodia; Leishmania; Trypanosoma; Schistosoma.
XX
OS Synthetic.
XX
PN WO9855495-A2.
XX
PD 10-DEC-1998.
XX
PF 05-JUN-1998; 98WO-US11578.
XX
PR 06-JUN-1997; 97US-0048793.
XX
PA (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI Dina D, Roman M, Schwartz D;
XX
DR WPI; 1999-059898/05.
XX
PT Immunostimulatory oligonucleotides regulate the immune system - and
PT contain an immune-stimulating octanucleotide sequence; for treating
PT cancer, allergic and infectious diseases
XX
PS Example 1; Page 29; 63pp; English.
XX
CC The invention relates to immunomodulatory oligonucleotides that comprise
CC at least 1 immunostimulatory octanucleotide sequence (ISS) where the ISS
CC sequences are selected from the group consisting of AACGTTCC, AACGTTCG,
CC GACGTTCC, and GACGTTCG. The immunomodulatory sequences are used to treat
CC patients needing immune regulation, such as those suffering from cancer,
CC an allergic disease and asthma. They are also used to prevent infectious
CC diseases such as influenza, herpes, hepatitis B, human immunodeficiency
CC and papillomavirus, Hemophilus influenza, Mycobacterium tuberculosis and
CC Bordetella pertussis, malarial plasmodia, Leishmania, Trypanosoma and
CC Schistosoma. The immunomodulatory sequences are used to screen for human
CC immunostimulatory activity by incubating macrophage cells and the
CC oligonucleotide; and determining the relative amount of Th1-biased
CC cytokines in the supernatant. Sequences AAV80104 to AAV80116 represent
CC oligonucleotides that were tested for immunostimulatory activity. These
CC were used in experiments for the stimulation of cytokine production and
CC were found to lack immunostimulatory activity. The invention provides
CC specific claimed examples (AAV80096-103) of immunomodulatory sequences.
XX
SQ Sequence 22 BP; 5 A; 7 C; 4 G; 6 T; 0 other;

Query Match 92.7%; Score 20.4; DB 20; Length 22;
Best Local Similarity 95.5%; Pred. No. 0.23;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 TGACTGTGAACGTTCGAGATGA 22
           |||| ||||||||||||||||||
Db      22 TGACCGTGAACGTTCGAGATGA 1
```

RESULT 27
AAV80096
ID AAV80096 standard; DNA; 22 BP.
XX
AC AAV80096;
XX
DT 12-MAR-1999 (first entry)
XX
DE Immunomodulatory oligo comprising an ISS sequence.
XX
KW Immunomodulatory; immunostimulatory; octanucleotide; immune regulation;
KW ISS: cancer; allergy; asthma; hepatitis B infection; papillomavirus;
KW human immunodeficiency virus; influenza; herpes; M. tuberculosis; ss;
KW B. pertussis; malaria; plasmodia; Leishmania; Trypanosoma; Schistosoma.
XX
OS Synthetic.
XX
PN WO9855495-A2.
XX
PD 10-DEC-1998.
XX
PF 05-JUN-1998; 98WO-US11578.
XX
PR 06-JUN-1997; 97US-0048793.
XX
PA (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI Dina D, Roman M, Schwartz D;
XX
DR WPI; 1999-059898/05.

```
XX
PT   Immunostimulatory oligonucleotides regulate the immune system - and
PT   contain an immune-stimulating octanucleotide sequence; for treating
PT   cancer, allergic and infectious diseases
XX
PS   Claim 7; Page 29; 63pp; English.
XX
CC   The invention relates to immunomodulatory oligonucleotides that comprise
CC   at least 1 immunostimulatory octanucleotide sequence (ISS) where the ISS
CC   sequences are selected from the group consisting of AACGTTCC, AACGTTCG,
CC   GACGTTCC, and GACGTTCG. The immunomodulatory sequences are used to treat
CC   patients needing immune regulation, such as those suffering from cancer,
CC   an allergic disease and asthma. They are also used to prevent infectious
CC   diseases such as influenza, herpes, hepatitis B, human immunodeficiency
CC   and papillomavirus, Hemophilus influenza, Mycobacterium tuberculosis and
CC   Bordetella pertussis, malarial plasmodia, Leishmania, Trypanosoma and
CC   Schistosoma. The immunomodulatory sequences are used to screen for human
CC   immunostimulatory activity by incubating macrophage cells and the
CC   oligonucleotide; and determining the relative amount of Th1-biased
CC   cytokines in the supernatant. Sequences AAV80096 to AAV80103 represent
CC   specific claimed examples of such immunomodulatory oligonucleotides.
XX
SQ   Sequence 22 BP; 6 A; 4 C; 7 G; 5 T; 0 other;

  Query Match             92.7%;  Score 20.4;  DB 20;  Length 22;
  Best Local Similarity   95.5%;  Pred. No. 0.23;
  Matches   21;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           |||| ||||||||||||||||||
Db       1 tgaccgtgaacgttcgagatga 22


RESULT  28
AAV80099
ID   AAV80099 standard; DNA; 22 BP.
XX
AC   AAV80099;
XX
DT   12-MAR-1999  (first entry)
XX
DE   Immunomodulatory oligo comprising an ISS sequence.
XX
KW   Immunomodulatory; immunostimulatory; octanucleotide; immune regulation;
KW   ISS: cancer; allergy; asthma; hepatitis B infection; papillomavirus;
KW   human immunodeficiency virus; influenza; herpes; M. tuberculosis; ss;
KW   B. pertussis; malaria; plasmodia; Leishmania; Trypanosoma; Schistosoma.
XX
OS   Synthetic.
XX
PN   WO9855495-A2.
XX
PD   10-DEC-1998.
XX
PF   05-JUN-1998;   98WO-US11578.
XX
PR   06-JUN-1997;   97US-0048793.
XX
PA   (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI   Dina D,  Roman M,  Schwartz D;
XX
DR   WPI; 1999-059898/05.
XX
PT   Immunostimulatory oligonucleotides regulate the immune system - and
PT   contain an immune-stimulating octanucleotide sequence; for treating
PT   cancer, allergic and infectious diseases
XX
PS   Claim 8; Page 29; 63pp; English.
XX
CC   The invention relates to immunomodulatory oligonucleotides that comprise
CC   at least 1 immunostimulatory octanucleotide sequence (ISS) where the ISS
CC   sequences are selected from the group consisting of AACGTTCC, AACGTTCG,
CC   GACGTTCC, and GACGTTCG. The immunomodulatory sequences are used to treat
CC   patients needing immune regulation, such as those suffering from cancer,
CC   an allergic disease and asthma. They are also used to prevent infectious
CC   diseases such as influenza, herpes, hepatitis B, human immunodeficiency
CC   and papillomavirus, Hemophilus influenza, Mycobacterium tuberculosis and
CC   Bordetella pertussis, malarial plasmodia, Leishmania, Trypanosoma and
CC   Schistosoma. The immunomodulatory sequences are used to screen for human
CC   immunostimulatory activity by incubating macrophage cells and the
CC   oligonucleotide; and determining the relative amount of Th1-biased
CC   cytokines in the supernatant. Sequences AAV80096 to AAV80103 represent
CC   specific claimed examples of such immunomodulatory oligonucleotides.
XX
SQ   Sequence 22 BP; 6 A; 4 C; 6 G; 6 T; 0 other;

  Query Match             92.7%;  Score 20.4;  DB 20;  Length 22;
  Best Local Similarity   95.5%;  Pred. No. 0.23;
  Matches   21;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           |||||||||||||||| ||||||
Db       1 tgactgtgaacgttccagatga 22


RESULT  29
AAV80101
ID   AAV80101 standard; DNA; 22 BP.
XX
AC   AAV80101;
XX
DT   12-MAR-1999  (first entry)
XX
DE   Immunomodulatory oligo comprising an ISS sequence.
XX
KW   Immunomodulatory; immunostimulatory; octanucleotide; immune regulation;
KW   ISS: cancer; allergy; asthma; hepatitis B infection; papillomavirus;
KW   human immunodeficiency virus; influenza; herpes; M. tuberculosis; ss;
KW   B. pertussis; malaria; plasmodia; Leishmania; Trypanosoma; Schistosoma.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   modified_base    11
FT                    /*tag= a
FT                    /note= "5-bromocytosine"
XX
PN   WO9855495-A2.
XX
PD   10-DEC-1998.
XX
PF   05-JUN-1998;   98WO-US11578.
XX
PR   06-JUN-1997;   97US-0048793.
XX
PA   (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI   Dina D,  Roman M,  Schwartz D;
XX
DR   WPI; 1999-059898/05.
XX
PT   Immunostimulatory oligonucleotides regulate the immune system - and
PT   contain an immune-stimulating octanucleotide sequence; for treating
PT   cancer, allergic and infectious diseases
XX
PS   Claim 22; Page 30; 63pp; English.
XX
CC   The invention relates to immunomodulatory oligonucleotides that comprise
CC   at least 1 immunostimulatory octanucleotide sequence (ISS) where the ISS
CC   sequences are selected from the group consisting of AACGTTCC, AACGTTCG,
CC   GACGTTCC, and GACGTTCG. The immunomodulatory sequences are used to treat
CC   patients needing immune regulation, such as those suffering from cancer,
CC   an allergic disease and asthma. They are also used to prevent infectious
```

```
CC   diseases such as influenza, herpes, hepatitis B, human immunodeficiency
CC   and papillomavirus, Hemophilus influenza, Mycobacterium tuberculosis and
CC   Bordetella pertussis, malarial plasmodia, Leishmania, Trypanosoma and
CC   Schistosoma. The immunomodulatory sequences are used to screen for human
CC   immunostimulatory activity by incubating macrophage cells and the
CC   oligonucleotide; and determining the relative amount of Th1-biased
CC   cytokines in the supernatant. Sequences AAV80096 to AAV80103 represent
CC   specific claimed examples of such immunomodulatory oligonucleotides.
XX
SQ   Sequence 22 BP; 6 A; 4 C; 6 G; 6 T; 0 other;

  Query Match             92.7%;  Score 20.4;  DB 20;  Length 22;
  Best Local Similarity   95.5%;  Pred. No. 0.23;
  Matches   21;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           |||||||||||||||| ||||||
Db       1 tgactgtgaacgttccagatga 22


RESULT  30
AAA96254
ID   AAA96254 standard; DNA; 22 BP.
XX
AC   AAA96254;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Sequence of a stabilised oligonucleotide with antitumour activity.
XX
KW   Antitumour; immunostimulatory oligonucleotide; tumour; anaplasia;
KW   glioblastoma; medullablastoma; neuroblastoma; melanoma; carcinoma; ss.
XX
OS   Synthetic.
XX
PN   WO200056342-A2.
XX
PD   28-SEP-2000.
XX
PF   17-MAR-2000; 2000WO-FR00676.
XX
PR   19-MAR-1999;   99FR-0003433.
XX
PA   (ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.
PA   (INRM ) INST NAT SANTE & RECH MEDICALE.
XX
PI   Carpentier A;
XX
DR   WPI; 2000-602192/57.
XX
PT   Use of stabilized oligonucleotides as antitumor agents, particularly
PT   against nervous system tumors, have optimal activity and are not toxic
PT   -
XX
PS   Example 13; Page 46; 57pp; French.
XX
CC   The present sequence represents a stabilised oligonucleotide which has
CC   antitumour activity. The oligonucleotide comprises an octomer motif
CC   of the type 5'-purine-purine-CG-pyrimidine-pyrimidine-X-X-3', where
CC   the pair X-X is AT, AA, CT or TT. The oligonucleotides are
CC   immunostimulatory, and are not toxic. They may be adapted for use in
CC   animals or humans. The stabilised oligonucleotides are used for
CC   treating tumours, of any type and any degree of anaplasia, particularly
CC   human tumours in the peripheral or central nervous systems, specifically
CC   glioblastomas, medullablastomas, neuroblastomas, melanomas or carcinomas.
XX
SQ   Sequence 22 BP; 6 A; 4 C; 6 G; 6 T; 0 other;

  Query Match             92.7%;  Score 20.4;  DB 21;  Length 22;
  Best Local Similarity   95.5%;  Pred. No. 0.23;
  Matches   21;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           |||||||||||||||| ||||||
Db       1 tgactgtgaacgttccagatga 22


RESULT  31
AAA38066
ID   AAA38066 standard; DNA; 22 BP.
XX
AC   AAA38066;
XX
DT   24-AUG-2000  (first entry)
XX
DE   Immunostimulatory sequence (ISS) #2.
XX
KW   Immunostimulatory sequence; ISS; immunomodulator; glycoprotein 120;
KW   gp120; human immunodeficiency virus; HIV; immune response; infection;
KW   development; ss.
XX
OS   Synthetic.
XX
PN   WO200021556-A1.
XX
PD   20-APR-2000.
XX
PF   08-OCT-1999;   99WO-US23677.
XX
PR   09-OCT-1998;   98US-0103733.
PR   07-OCT-1999;   99US-0415186.
XX
PA   (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI   Tighe H,  Raz E,  Schwartz D,  Takabayashi K;
XX
DR   WPI; 2000-317846/27.
XX
PT   Anti-HIV composition comprises immunostimulatory polynucleotides and
PT   HIV glycoprotein gp120 useful for modulating, stimulating an immune
PT   response against HIV in an HIV infected individual  -
XX
PS   Disclosure; Page 16; 65pp; English.
XX
CC   The present invention relates to an immunostimulatory composition
CC   comprising a human immunodeficiency virus (HIV) antigen, and an
CC   immunomodulatory polynucleotide comprising an immunostimulatory sequence
CC   (ISS). This sequence represents an ISS that can be used in the
CC   composition. An immunostimulatory composition which comprises a gp120
CC   conjugated to an immunomodulatory polynucleotide, or is proximately
CC   associated to it and not conjugated, is used for modulating or
CC   stimulating a specific immune response against gp120 in an individual by
CC   producing anti-gp120 antibodies or gp120 specific cytotoxic T cells. It
CC   is also used for suppressing or delaying development of HIV infection in
CC   an individual infected with HIV or an individual at risk of infection
CC   with HIV, respectively. It is also used for treating an individual
CC   infected with HIV in need of immune modulation.
XX
SQ   Sequence 22 BP; 6 A; 4 C; 7 G; 5 T; 0 other;

  Query Match             92.7%;  Score 20.4;  DB 21;  Length 22;
  Best Local Similarity   95.5%;  Pred. No. 0.23;
  Matches   21;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           |||| ||||||||||||||||||
Db       1 tgaccgtgaacgttcgagatga 22


RESULT  32
AAA38068
ID   AAA38068 standard; DNA; 22 BP.
XX
```

```
AC   AAA38068;
XX
DT   24-AUG-2000  (first entry)
XX
DE   Immunostimulatory sequence (ISS) #4.
XX
KW   Immunostimulatory sequence; ISS; immunomodulator; glycoprotein 120;
KW   gp120; human immunodeficiency virus; HIV; immune response; infection;
KW   development; ss.
XX
OS   Synthetic.
XX
PN   WO200021556-A1.
XX
PD   20-APR-2000.
XX
PF   08-OCT-1999;   99WO-US23677.
XX
PR   09-OCT-1998;   98US-0103733.
PR   07-OCT-1999;   99US-0415186.
XX
PA   (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI   Tighe H,  Raz E,  Schwartz D,  Takabayashi K;
XX
DR   WPI; 2000-317846/27.
XX
PT   Anti-HIV composition comprises immunostimulatory polynucleotides and
PT   HIV glycoprotein gp120 useful for modulating, stimulating an immune
PT   response against HIV in an HIV infected individual  -
XX
PS   Disclosure; Page 16; 65pp; English.
XX
CC   The present invention relates to an immunostimulatory composition
CC   comprising a human immunodeficiency virus (HIV) antigen, and an
CC   immunomodulatory polynucleotide comprising an immunostimulatory sequence
CC   (ISS). This sequence represents an ISS that can be used in the
CC   composition. An immunostimulatory composition which comprises a gp120
CC   conjugated to an immunomodulatory polynucleotide, or is proximately
CC   associated to it and not conjugated, is used for modulating or
CC   stimulating a specific immune response against gp120 in an individual by
CC   producing anti-gp120 antibodies or gp120 specific cytotoxic T cells. It
CC   is also used for suppressing or delaying development of HIV infection in
CC   an individual infected with HIV or an individual at risk of infection
CC   with HIV, respectively. It is also used for treating an individual
CC   infected with HIV in need of immune modulation.
XX
SQ   Sequence 22 BP; 6 A; 4 C; 6 G; 6 T; 0 other;
```

```
  Query Match             92.7%;  Score 20.4;  DB 21;  Length 22;
  Best Local Similarity   95.5%;  Pred. No. 0.23;
  Matches   21;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           |||||||||||||| ||||||
Db       1 tgactgtgaacgttccagatga 22
```

```
RESULT  33
AAA38070
ID   AAA38070 standard; DNA; 22 BP.
XX
AC   AAA38070;
XX
DT   24-AUG-2000  (first entry)
XX
DE   Immunostimulatory sequence (ISS) #6.
XX
KW   Immunostimulatory sequence; ISS; immunomodulator; glycoprotein 120;
KW   gp120; human immunodeficiency virus; HIV; immune response; infection;
KW   development; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   modified_base   11
FT                   /*tag= a
FT                   /mod_base= OTHER
FT                   /note= "5-Bromocytosine"
XX
PN   WO200021556-A1.
XX
PD   20-APR-2000.
XX
PF   08-OCT-1999;   99WO-US23677.
XX
PR   09-OCT-1998;   98US-0103733.
PR   07-OCT-1999;   99US-0415186.
XX
PA   (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI   Tighe H,  Raz E,  Schwartz D,  Takabayashi K;
XX
DR   WPI; 2000-317846/27.
XX
PT   Anti-HIV composition comprises immunostimulatory polynucleotides and
PT   HIV glycoprotein gp120 useful for modulating, stimulating an immune
PT   response against HIV in an HIV infected individual  -
XX
PS   Disclosure; Page 16; 65pp; English.
XX
CC   The present invention relates to an immunostimulatory composition
CC   comprising a human immunodeficiency virus (HIV) antigen, and an
CC   immunomodulatory polynucleotide comprising an immunostimulatory sequence
CC   (ISS). This sequence represents an ISS that can be used in the
CC   composition. An immunostimulatory composition which comprises a gp120
CC   conjugated to an immunomodulatory polynucleotide, or is proximately
CC   associated to it and not conjugated, is used for modulating or
CC   stimulating a specific immune response against gp120 in an individual by
CC   producing anti-gp120 antibodies or gp120 specific cytotoxic T cells. It
CC   is also used for suppressing or delaying development of HIV infection in
CC   an individual infected with HIV or an individual at risk of infection
CC   with HIV, respectively. It is also used for treating an individual
CC   infected with HIV in need of immune modulation.
XX
SQ   Sequence 22 BP; 6 A; 4 C; 6 G; 6 T; 0 other;
```

```
  Query Match             92.7%;  Score 20.4;  DB 21;  Length 22;
  Best Local Similarity   95.5%;  Pred. No. 0.23;
  Matches   21;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           |||||||||||||| ||||||
Db       1 tgactgtgaacgttccagatga 22
```

```
RESULT  34
AAH42534
ID   AAH42534 standard; DNA; 22 BP.
XX
AC   AAH42534;
XX
DT   01-OCT-2001  (first entry)
XX
DE   Phosphorothioate beta-gal/immunostimulatory mutated oligonucleotide.
XX
KW   Anaphylactic hypersensitivity; immunomodulatory nucleic acid; vaccine;
KW   anaphylaxis-associated symptom; IgE; histamine; phosphorothioate; ss.
XX
OS   Synthetic.
XX
PN   WO200145750-A1.
XX
PD   28-JUN-2001.
```

```
XX
PF   20-DEC-2000; 2000WO-US35064.
XX
PR   21-DEC-1999;   99US-0171830.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Raz E,  Horner AA;
XX
DR   WPI; 2001-475812/51.
XX
PT   Reducing risk of anaphylactic hypersensitivity response to an allergen
PT   in a subject, by administering an immunomodulating nucleic acid
PT   molecule comprising a specific sequence  -
XX
PS   Example 1; Page 23; 39pp; English.
XX
CC   The specification describes a method for reducing a symptom associated
CC   with anaphylactic hypersensitivity or risk of anaphylactic response in
CC   a subject. The method comprises administering to an individual a
CC   nucleic acid molecule comprising an immunomodulatory nucleic acid
CC   molecule (INA) comprising the sequence 5'-C-G-3' to reduce
CC   anaphylaxis-associated symptom. The method is useful for reducing a
CC   symptom associated with anaphylactic hypersensitivity, including
CC   elevated IgE level, elevated histamine level, constriction of the
CC   airways and difficult breathing which can lead to anaphylactic reaction
CC   or anaphylactic shock, thereby reducing the risk of death. The present
CC   sequence represents a beta-gal/immunostimulatory mutated sequence, which
CC   was used as a vaccine to protect against the development of anaphylactic
CC   hypersensitivity.
XX
SQ   Sequence 22 BP; 6 A; 2 C; 8 G; 6 T; 0 other;

  Query Match             92.7%;  Score 20.4;  DB 22;  Length 22;
  Best Local Similarity   95.5%;  Pred. No. 0.23;
  Matches   21;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           |||||||||| ||||||||||||
Db       1 tgactgtgaaggttcgagatga 22


RESULT  35
AAH73440
ID   AAH73440 standard; DNA; 22 BP.
XX
AC   AAH73440;
XX
DT   01-OCT-2001  (first entry)
XX
DE   Immunomodulatory nucleic acid control sequence #1.
XX
KW   G3PDH gene; immunomodulatory oligonucleotide; infection; mycobacterium;
KW   intracellular pathogen; anti-pathogenic; ss.
XX
OS   Unidentified.
XX
PN   WO200155341-A2.
XX
PD   02-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US03029.
XX
PR   31-JAN-2000; 2000US-0179353.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Raz E,  Kornbluth R,  Catanzaro A,  Hayashi T,  Carson DA;
XX
DR   WPI; 2001-483234/52.
XX
PT   Treating infection of intracellular pathogen e.g., Mycobacterium, in a
PT   subject, involves administering immunomodulatory nucleic acid molecule
PT   to inhibit intracellular replication of intracellular pathogen  -
XX
PS   Disclosure; Page 13; 54pp; English.
XX
CC   The present invention describes a method of treating an infection caused
CC   by an intracellular pathogen, involving administering to the patient an
CC   immunomodulatory nucleic acid and an anti-pathogenic agent. This is
CC   particularly useful in the treatment of mycobacterial infections. The
CC   present sequence is a control sequence to an immunomodulatory nucleic
CC   acid described in the exemplification of the invention.
XX
SQ   Sequence 22 BP; 6 A; 2 C; 8 G; 6 T; 0 other;

  Query Match             92.7%;  Score 20.4;  DB 22;  Length 22;
  Best Local Similarity   95.5%;  Pred. No. 0.23;
  Matches   21;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           |||||||||| ||||||||||||
Db       1 tgactgtgaaggttcgagatga 22


RESULT  36
AAH41574
ID   AAH41574 standard; DNA; 22 BP.
XX
AC   AAH41574;
XX
DT   24-AUG-2001  (first entry)
XX
DE   Immunostimulatory sequence (ISS) SEQ ID NO:2.
XX
KW   Immunostimulatory sequence; ISS; immunomodulatory; immune response;
KW   antigen; antiallergic; modulation; Th1 lymphocyte stimulation; allergy;
KW   Th1-associated cytokine; Th2 lymphocyte suppression; cytokine; ss.
XX
OS   Synthetic.
XX
PN   WO200135991-A2.
XX
PD   25-MAY-2001.
XX
PF   15-NOV-2000; 2000WO-US31385.
XX
PR   15-NOV-1999;   99US-0165467.
PR   14-NOV-2000; 2000US-0713136.
XX
PA   (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI   Tuck S,  Van Nest G;
XX
DR   WPI; 2001-329209/34.
XX
PT   Populations of conjugate molecules comprising polynucleotide
PT   immunostimulatory sequences polynucleotides and antigens, useful for
PT   controlling immune responses  -
XX
PS   Disclosure; Page 30; 97pp; English.
XX
CC   The present invention describes immunomodulatory populations ((I) and
CC   (II)) of conjugate molecules (CMs) comprising immunostimulatory sequences
CC   (ISS) of polynucleotides and antigens. The extent of conjugation affects
CC   the immunological properties (e.g. the extent of antigen-specific
CC   antibody formation, including Th1-associated antibody formation) so the
CC   conjugates are used for altering the type and extent of immune response.
CC   (I) and (II) have immunomodulatory, immunosuppressive and antiallergic
CC   activities, and can be used in the modulation of immune responses via
CC   the stimulation of Th1 lymphocytes and Th1-associated cytokines, and
CC   suppression of Th2 lymphocytes and cytokines. The populations ((I) and
CC   (II)) of conjugate molecules may be used for modulating immune responses
CC   in individuals e.g. for the treatment of an allergic condition. (I) and
```

```
CC   (II) may be used to modulate immune responses and therefore prevent
CC   potentially harmful reactions to antigens. The present sequence
CC   represents an ISS polynucleotide which is used in the exemplification
CC   of the present invention.
XX
SQ   Sequence 22 BP; 6 A; 4 C; 7 G; 5 T; 0 other;


  Query Match             92.7%;  Score 20.4;  DB 22;  Length 22;
  Best Local Similarity   95.5%;  Pred. No. 0.23;
  Matches   21;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           |||| ||||||||||||||||||
Db       1 tgaccgtgaacgttcgagatga 22


RESULT  37
AAH41576
ID   AAH41576 standard; DNA; 22 BP.
XX
AC   AAH41576;
XX
DT   24-AUG-2001  (first entry)
XX
DE   Immunostimulatory sequence (ISS) SEQ ID NO:4.
XX
KW   Immunostimulatory sequence; ISS; immunomodulatory; immune response;
KW   antigen; antiallergic; modulation; Th1 lymphocyte stimulation; allergy;
KW   Th1-associated cytokine; Th2 lymphocyte suppression; cytokine; ss.
XX
OS   Synthetic.
XX
PN   WO200135991-A2.
XX
PD   25-MAY-2001.
XX
PF   15-NOV-2000; 2000WO-US31385.
XX
PR   15-NOV-1999;   99US-0165467.
PR   14-NOV-2000; 2000US-0713136.
XX
PA   (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI   Tuck S,  Van Nest G;
XX
DR   WPI; 2001-329209/34.
XX
PT   Populations of conjugate molecules comprising polynucleotide
PT   immunostimulatory sequences polynucleotides and antigens, useful for
PT   controlling immune responses -
XX
PS   Disclosure; Page 30; 97pp; English.
XX
CC   The present invention describes immunomodulatory populations ((I) and
CC   (II)) of conjugate molecules (CMs) comprising immunostimulatory sequences
CC   (ISS) of polynucleotides and antigens. The extent of conjugation affects
CC   the immunological properties (e.g. the extent of antigen-specific
CC   antibody formation, including Th1-associated antibody formation) so the
CC   conjugates are used for altering the type and extent of immune response.
CC   (I) and (II) have immunomodulatory, immunosuppressive and antiallergic
CC   activities, and can be used in the modulation of immune responses via
CC   the stimulation of Th1 lymphocytes and Th1-associated cytokines, and
CC   suppression of Th2 lymphocytes and cytokines. The populations ((I) and
CC   (II)) of conjugate molecules may be used for modulating immune responses
CC   in individuals e.g. for the treatment of an allergic condition. (I) and
CC   (II) may be used to modulate immune responses and therefore prevent
CC   potentially harmful reactions to antigens. The present sequence
CC   represents an ISS polynucleotide which is used in the exemplification
CC   of the present invention.
XX
SQ   Sequence 22 BP; 6 A; 4 C; 6 G; 6 T; 0 other;


  Query Match             92.7%;  Score 20.4;  DB 22;  Length 22;
  Best Local Similarity   95.5%;  Pred. No. 0.23;
  Matches   21;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           |||||||||||||||| ||||||
Db       1 tgactgtgaacgttccagatga 22


RESULT  38
AAF77041
ID   AAF77041 standard; DNA; 22 BP.
XX
AC   AAF77041;
XX
DT   15-MAY-2001  (first entry)
XX
DE   Immunostimulatory DNA #1.
XX
KW   Modulate; immune; antigen; immunostimulatory; ds.
XX
OS   Synthetic.
XX
PN   WO200112223-A2.
XX
PD   22-FEB-2001.
XX
PF   18-AUG-2000; 2000WO-US22835.
XX
PR   19-AUG-1999;   99US-0149768.
XX
PA   (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI   Van Nest G;
XX
DR   WPI; 2001-211136/21.
XX
PT   Modulating immune response to a second antigen in humans involves
PT   administering an immunostimulatory polynucleotide comprising an
PT   immunostimulatory sequence and a first antigen -
XX
PS   Disclosure; Page 15; 63pp; English.
XX
CC   The present invention relates to modulating an immune response to
CC   a second antigen in an individual, involving
CC   administering to the individual an immunomodulatory polynucleotide
CC   comprising an immunostimulatory sequence (ISS) and a first antigen.
XX
SQ   Sequence 22 BP; 6 A; 4 C; 7 G; 5 T; 0 other;


  Query Match             92.7%;  Score 20.4;  DB 22;  Length 22;
  Best Local Similarity   95.5%;  Pred. No. 0.23;
  Matches   21;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           |||| ||||||||||||||||||
Db       1 tgaccgtgaacgttcgagatga 22


RESULT  39
AAF77043
ID   AAF77043 standard; DNA; 22 BP.
XX
AC   AAF77043;
XX
DT   15-MAY-2001  (first entry)
XX
DE   Immunostimulatory DNA #3.
XX
KW   Modulate; immune; antigen; immunostimulatory; ds.
XX
```

OS Synthetic.
XX
PN WO200112223-A2.
XX
PD 22-FEB-2001.
XX
PF 18-AUG-2000; 2000WO-US22835.
XX
PR 19-AUG-1999; 99US-0149768.
XX
PA (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI Van Nest G;
XX
DR WPI; 2001-211136/21.
XX
PT Modulating immune response to a second antigen in humans involves
PT administering an immunostimulatory polynucleotide comprising an
PT immunostimulatory sequence and a first antigen -
XX
PS Disclosure; Page 15; 63pp; English.
XX
CC The present invention relates to modulating an immune response to
CC a second antigen in an individual, involving
CC administering to the individual an immunomodulatory polynucleotide
CC comprising an immunostimulatory sequence (ISS) and a first antigen.
XX
SQ Sequence 22 BP; 6 A; 4 C; 6 G; 6 T; 0 other;

Query Match 92.7%; Score 20.4; DB 22; Length 22;
Best Local Similarity 95.5%; Pred. No. 0.23;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TGACTGTGAACGTTCGAGATGA 22
||||||||||||||| ||||||
Db 1 tgactgtgaacgttccagatga 22

RESULT 40
AAF77047
ID AAF77047 standard; DNA; 22 BP.
XX
AC AAF77047;
XX
DT 15-MAY-2001 (first entry)
XX
DE Immunostimulatory DNA #7.
XX
KW Modulate; immune; antigen; immunostimulatory; ds.
XX
OS Synthetic.
XX
PN WO200112223-A2.
XX
PD 22-FEB-2001.
XX
PF 18-AUG-2000; 2000WO-US22835.
XX
PR 19-AUG-1999; 99US-0149768.
XX
PA (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI Van Nest G;
XX
DR WPI; 2001-211136/21.
XX
PT Modulating immune response to a second antigen in humans involves
PT administering an immunostimulatory polynucleotide comprising an
PT immunostimulatory sequence and a first antigen -
XX
PS Disclosure; Page 15; 63pp; English.
XX
CC The present invention relates to modulating an immune response to
CC a second antigen in an individual, involving
CC administering to the individual an immunomodulatory polynucleotide
CC comprising an immunostimulatory sequence (ISS) and a first antigen.
XX
SQ Sequence 22 BP; 6 A; 1 C; 7 G; 6 T; 2 other;

Query Match 92.7%; Score 20.4; DB 22; Length 22;
Best Local Similarity 90.9%; Pred. No. 0.23;
Matches 20; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGACTGTGAACGTTCGAGATGA 22
|||||||||||:|||:|||||||
Db 1 tgactgtgaabgttbgagatga 22

RESULT 41
AAH41580
ID AAH41580 standard; DNA; 22 BP.
XX
AC AAH41580;
XX
DT 24-AUG-2001 (first entry)
XX
DE Immunostimulatory sequence (ISS) SEQ ID NO:8.
XX
KW Immunostimulatory sequence; ISS; immunomodulatory; immune response;
KW antigen; antiallergic; modulation; Th1 lymphocyte stimulation; allergy;
KW Th1-associated cytokine; Th2 lymphocyte suppression; cytokine; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT modified_base 11
FT /*tag= a
FT /mod_base= "OTHER"
FT /note= "5-bromocytosine"
XX
PN WO200135991-A2.
XX
PD 25-MAY-2001.
XX
PF 15-NOV-2000; 2000WO-US31385.
XX
PR 15-NOV-1999; 99US-0165467.
PR 14-NOV-2000; 2000US-0713136.
XX
PA (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI Tuck S, Van Nest G;
XX
DR WPI; 2001-329209/34.
XX
PT Populations of conjugate molecules comprising polynucleotide
PT immunostimulatory sequences polynucleotides and antigens, useful for
PT controlling immune responses -
XX
PS Disclosure; Page 31; 97pp; English.
XX
CC The present invention describes immunomodulatory populations ((I) and
CC (II)) of conjugate molecules (CMs) comprising immunostimulatory sequences
CC (ISS) of polynucleotides and antigens. The extent of conjugation affects
CC the immunological properties (e.g. the extent of antigen-specific
CC antibody formation, including Th1-associated antibody formation) so the
CC conjugates are used for altering the type and extent of immune response.
CC (I) and (II) have immunomodulatory, immunosuppressive and antiallergic
CC activities, and can be used in the modulation of immune responses via
CC the stimulation of Th1 lymphocytes and Th1-associated cytokines, and
CC suppression of Th2 lymphocytes and cytokines. The populations ((I) and
CC (II)) of conjugate molecules may be used for modulating immune responses
CC in individuals e.g. for the treatment of an allergic condition. (I) and
CC (II) may be used to modulate immune responses and therefore prevent

```
CC   potentially harmful reactions to antigens. The present sequence
CC   represents an ISS polynucleotide which is used in the exemplification
CC   of the present invention.
XX
SQ   Sequence 22 BP; 6 A; 1 C; 7 G; 6 T; 2 other;


  Query Match           91.8%;  Score 20.2;  DB 22;  Length 22;
  Best Local Similarity  90.9%;  Pred. No. 0.29;
  Matches   20;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          |||||||||| ||| :|||||||
Db      1 tgactgtgaangttbgagatga 22


RESULT  42
AAZ55881
ID   AAZ55881 standard; DNA; 22 BP.
XX
AC   AAZ55881;
XX
DT   10-APR-2000  (first entry)
XX
DE   Immunomodulatory oligonucleotide SEQ ID NO: 6.
XX
KW   Immunomodulation; immunostimulatory sequence; adjuvant;
KW   Th1 immune response; cytotoxic T-cell; cytokine; cancer; allergy;
KW   asthma; immunocontraception; 5-bromocytosine; ss.
XX
OS   Mus musculus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   modified_base    1..22
FT                    /*tag= a
FT                    /note= "Phosphorothioate linkages"
FT   misc_feature     9..16
FT                    /*tag= b
FT                    /note= "Immunostimulatory sequence (ISS)"
FT   modified_base    11
FT                    /*tag= c
FT                    /mod_base= OTHER
FT                    /note= "5-bromocytosine"
FT   modified_base    15
FT                    /*tag= d
FT                    /mod_base= OTHER
FT                    /note= "5-bromocytosine"
XX
PN   WO9962923-A2.
XX
PD   09-DEC-1999.
XX
PF   04-JUN-1999;   99WO-US12538.
XX
PR   05-JUN-1998;   98US-0088310.
PR   01-JUN-1999;   99US-0324191.
XX
PA   (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI   Schwartz D;
XX
DR   WPI; 2000-105687/09.
XX
PT   Novel immunomodulatory oligonucleotide used to induce a Th1-type immune
PT   response, e.g. to tumor antigens  -
XX
PS   Claim 31; Page 35; 54pp; English.
XX
CC   Sequences AAZ55876-Z55877 and AAZ55880-Z55886 represent immunomodulatory
CC   oligonucleotides comprising an immunostimulatory sequence (ISS, e.g.,
CC   AACGTC, AACGTT, AGCGTC, AGCGCT, AGCGTT, GACGTC, GACGTT, GGCGTT,
CC   AACGTTCC and GACGTTCC). The invention relates to oligonucleotides
CC   comprising one or more ISSs, where the ISS comprises at least
CC   one modified cytosine with an electron-withdrawing molety at
CC   position C-5 or C-6 of the base. Sequences AAZ55877 and AAZ55880-Z55886
CC   contain ISSs comprising at least one bromocytosine, whereas sequence
CC   AAZ55876 contains an unmodified ISS. The immunomodulatory
CC   oligonucleotides have an adjuvant-like effect; when formulated with an
CC   antigen, the oligonucleotides stimulate production of Th1-type cytokines,
CC   and induce a Th1-type immune response (activation of cytotoxic T cells),
CC   while simultaneously downregulating the Th2-type response. The Th1
CC   response is particularly effective for control of viruses and
CC   intracellular parasites. The immunomodulatory oligonucleotides are used,
CC   particularly when formulated with an antigen or a facilitator, for
CC   modulating immune responses. Such compositions may be used in tumour
CC   therapy, in treatment of allergy (including asthma), for inducing a
CC   vigorous cellular response (against a virus, bacterium, fungus or
CC   protozoan), and also in contraceptive vaccines based on sperm antigens.
XX
SQ   Sequence 22 BP; 6 A; 1 C; 7 G; 6 T; 2 other;


  Query Match           90.9%;  Score 20;  DB 21;  Length 22;
  Best Local Similarity  90.9%;  Pred. No. 0.36;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          |||||||||| ||| |||||||
Db      1 tgactgtgaangttngagatga 22


RESULT  43
AAF77045
ID   AAF77045 standard; DNA; 22 BP.
XX
AC   AAF77045;
XX
DT   15-MAY-2001  (first entry)
XX
DE   Immunostimulatory DNA #5.
XX
KW   Modulate; immune; antigen; immunostimulatory; ds.
XX
OS   Synthetic.
XX
PN   WO200112223-A2.
XX
PD   22-FEB-2001.
XX
PF   18-AUG-2000; 2000WO-US22835.
XX
PR   19-AUG-1999;   99US-0149768.
XX
PA   (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI   Van Nest G;
XX
DR   WPI; 2001-211136/21.
XX
PT   Modulating immune response to a second antigen in humans involves
PT   administering an immunostimulatory polynucleotide comprising an
PT   immunostimulatory sequence and a first antigen  -
XX
PS   Disclosure; Page 15; 63pp; English.
XX
CC   The present invention relates to modulating an immune response to
CC   a second antigen in an individual, involving
CC   administering to the individual an immunomodulatory polynucleotide
CC   comprising an immunostimulatory sequence (ISS) and a first antigen.
XX
SQ   Sequence 22 BP; 6 A; 3 C; 6 G; 6 T; 1 other;


  Query Match           89.1%;  Score 19.6;  DB 22;  Length 22;
  Best Local Similarity  90.9%;  Pred. No. 0.58;
```

```
  Matches   20;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           ||||||||||:|||| ||||||
Db       1 tgactgtgaabgttccagatga 22


RESULT  44
AAZ55877
ID   AAZ55877 standard; DNA; 22 BP.
XX
AC   AAZ55877;
XX
DT   10-APR-2000  (first entry)
XX
DE   Immunomodulatory oligonucleotide SEQ ID NO: 2.
XX
KW   Immunomodulation; immunostimulatory sequence; adjuvant;
KW   Th1 immune response; cytotoxic T-cell; cytokine; cancer; allergy;
KW   asthma; immunocontraception; 5-bromocytosine; ss.
XX
OS   Mus musculus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   modified_base   1..22
FT                   /*tag=  a
FT                   /note= "Phosphorothioate linkages"
FT   misc_feature    9..16
FT                   /*tag=  b
FT                   /note= "Immunostimulatory sequence (ISS)"
FT   modified_base   11
FT                   /*tag=  c
FT                   /mod_base= OTHER
FT                   /note= "5-bromocytosine"
XX
PN   WO9962923-A2.
XX
PD   09-DEC-1999.
XX
PF   04-JUN-1999;   99WO-US12538.
XX
PR   05-JUN-1998;   98US-0088310.
PR   01-JUN-1999;   99US-0324191.
XX
PA   (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI   Schwartz D;
XX
DR   WPI; 2000-105687/09.
XX
PT   Novel immunomodulatory oligonucleotide used to induce a Th1-type immune
PT   response, e.g. to tumor antigens -
XX
PS   Claim 29; Page 35; 54pp; English.
XX
CC   Sequences AAZ55876-Z55877 and AAZ55880-Z55886 represent immunomodulatory
CC   oligonucleotides comprising an immunostimulatory sequence (ISS, e.g.,
CC   AACGTC, AACGTT, AGCGTC, AGCGCT, AGCGTT, GACGTC, GACGTT, GGCGTT,
CC   AACGTTCC and GACGTTCC). The invention relates to oligonucleotides
CC   comprising one or more ISSs, where the ISS comprises at least
CC   one modified cytosine with an electron-withdrawing moiety at
CC   position C-5 or C-6 of the base. Sequences AAZ55877 and AAZ55880-Z55886
CC   contain ISSs comprising at least one bromocytosine, whereas sequence
CC   AAZ55876 contains an unmodified ISS. The immunomodulatory
CC   oligonucleotides have an adjuvant-like effect; when formulated with an
CC   antigen, the oligonucleotides stimulate production of Th1-type cytokines,
CC   and induce a Th1-type immune response (activation of cytotoxic T cells),
CC   while simultaneously downregulating the Th2-type response. The Th1
CC   response is particularly effective for control of viruses and
CC   intracellular parasites. The immunomodulatory oligonucleotides are used,
CC   particularly when formulated with an antigen or a facilitator, for
CC   modulating immune responses. Such compositions may be used in tumour
CC   therapy, in treatment of allergy (including asthma), for inducing a
CC   vigorous cellular response (against a virus, bacterium, fungus or
CC   protozoan), and also in contraceptive vaccines based on sperm antigens.
CC
XX
SQ   Sequence 22 BP; 6 A; 3 C; 6 G; 6 T; 1 other;


  Query Match             88.2%;  Score 19.4;  DB 21;  Length 22;
  Best Local Similarity   90.9%;  Pred. No. 0.73;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           |||||||||| |||| ||||||
Db       1 tgactgtgaangttccagatga 22


RESULT  45
AAH41578
ID   AAH41578 standard; DNA; 22 BP.
XX
AC   AAH41578;
XX
DT   24-AUG-2001  (first entry)
XX
DE   Immunostimulatory sequence (ISS) SEQ ID NO:6.
XX
KW   Immunostimulatory sequence; ISS; immunomodulatory; immune response;
KW   antigen; antiallergic; modulation; Th1 lymphocyte stimulation; allergy;
KW   Th1-associated cytokine; Th2 lymphocyte suppression; cytokine; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   modified_base   11
FT                   /*tag=  a
FT                   /mod_base= "OTHER"
FT                   /note= "5-bromocytosine"
XX
PN   WO200135991-A2.
XX
PD   25-MAY-2001.
XX
PF   15-NOV-2000; 2000WO-US31385.
XX
PR   15-NOV-1999;   99US-0165467.
PR   14-NOV-2000; 2000US-0713136.
XX
PA   (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI   Tuck S,  Van Nest G;
XX
DR   WPI; 2001-329209/34.
XX
PT   Populations of conjugate molecules comprising polynucleotide
PT   immunostimulatory sequences polynucleotides and antigens, useful for
PT   controlling immune responses -
XX
PS   Disclosure; Page 30; 97pp; English.
XX
CC   The present invention describes immunomodulatory populations ((I) and
CC   (II)) of conjugate molecules (CMs) comprising immunostimulatory sequences
CC   (ISS) of polynucleotides and antigens. The extent of conjugation affects
CC   the immunological properties (e.g. the extent of antigen-specific
CC   antibody formation, including Th1-associated antibody formation) so the
CC   conjugates are used for altering the type and extent of immune response.
CC   (I) and (II) have immunomodulatory, immunosuppressive and antiallergic
CC   activities, and can be used in the modulation of immune responses via
CC   the stimulation of Th1 lymphocytes and Th1-associated cytokines, and
CC   suppression of Th2 lymphocytes and cytokines. The populations ((I) and
CC   (II)) of conjugate molecules may be used for modulating immune responses
CC   in individuals e.g. for the treatment of an allergic condition. (I) and
CC   (II) may be used to modulate immune responses and therefore prevent
CC   potentially harmful reactions to antigens. The present sequence
```

CC represents an ISS polynucleotide which is used in the exemplification
CC of the present invention.
XX
SQ Sequence 22 BP; 6 A; 3 C; 6 G; 6 T; 1 other;

Query Match 88.2%; Score 19.4; DB 22; Length 22;
Best Local Similarity 90.9%; Pred. No. 0.73;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 TGACTGTGAACGTTCGAGATGA 22
           |||||||||| |||| ||||||
Db       1 tgactgtgaangttccagatga 22
```

Search completed: November 29, 2001, 14:51:06
Job time: 3659 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: November 29, 2001, 14:48:18 ; Search time 64.43 Seconds
(without alignments)
77.332 Million cell updates/sec

Title: SEQ1
Perfect score: 22
Sequence: 1 TGACTGTGAACGTTCGAGATGA 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 560984

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20.4	92.7	22	4	US-09-092-314-2	Sequence 2, Appli
	2	18.8	85.5	22	4	US-09-092-314-1	Sequence 1, Appli
	3	18.8	85.5	22	4	US-09-092-314-3	Sequence 3, Appli
	4	18.8	85.5	22	4	US-09-092-314-10	Sequence 10, Appl
	5	17.2	78.2	22	4	US-09-092-314-4	Sequence 4, Appli
	6	15.6	70.9	22	4	US-09-092-314-5	Sequence 5, Appli
	7	15.6	70.9	22	4	US-09-092-314-7	Sequence 7, Appli
	8	15.6	70.9	22	4	US-09-092-314-8	Sequence 8, Appli
	9	14	63.6	77	1	US-08-399-412A-58	Sequence 58, Appl
	10	14	63.6	95	4	US-08-952-793-258	Sequence 258, App
	11	14	63.6	95	5	PCT-US96-09455A-258	Sequence 258, App
	12	13.6	61.8	77	1	US-08-384-708A-195	Sequence 195, App
	13	13.6	61.8	77	4	US-08-687-421-287	Sequence 287, App
c	14	13.6	61.8	97	1	US-08-210-222-11	Sequence 11, Appl
c	15	13.6	61.8	98	1	US-08-210-222-8	Sequence 8, Appli
c	16	13.6	61.8	98	1	US-08-210-222-19	Sequence 19, Appl
c	17	13.6	61.8	98	1	US-08-210-222-22	Sequence 22, Appl
c	18	13.6	61.8	98	1	US-08-210-222-24	Sequence 24, Appl
c	19	13.4	60.9	71	3	US-08-633-768A-12	Sequence 12, Appl
	20	13.2	60.0	36	4	US-09-386-607-6	Sequence 6, Appli
	21	13.2	60.0	77	1	US-08-400-440A-19	Sequence 19, Appl
	22	13.2	60.0	77	1	US-08-463-093A-19	Sequence 19, Appl
	23	13.2	60.0	77	2	US-08-460-888A-19	Sequence 19, Appl
	24	13.2	60.0	77	2	US-08-894-578-19	Sequence 19, Appl
	25	13.2	60.0	77	4	US-09-412-017-19	Sequence 19, Appl
c	26	13.2	60.0	97	1	US-08-210-222-4	Sequence 4, Appli
c	27	13.2	60.0	98	1	US-08-210-222-7	Sequence 7, Appli
c	28	13	59.1	60	4	US-09-017-612A-1	Sequence 1, Appli
	29	12.8	58.2	31	4	US-09-070-408-110	Sequence 110, App
	30	12.8	58.2	36	1	US-08-403-762A-163	Sequence 163, App
	31	12.8	58.2	77	1	US-08-447-169A-17	Sequence 17, Appl
	32	12.8	58.2	77	2	US-08-233-012C-17	Sequence 17, Appl
	33	12.6	57.3	26	1	US-08-403-762A-149	Sequence 149, App
	34	12.6	57.3	27	2	US-08-308-952-18	Sequence 18, Appl
	35	12.6	57.3	27	4	US-09-124-141-27	Sequence 27, Appl
	36	12.6	57.3	36	1	US-08-153-799-12	Sequence 12, Appl
c	37	12.6	57.3	57	4	US-09-017-612A-3	Sequence 3, Appli
c	38	12.6	57.3	59	1	US-08-440-084-7	Sequence 7, Appli
c	39	12.6	57.3	59	5	PCT-US96-06669-7	Sequence 7, Appli
	40	12.6	57.3	61	1	US-07-744-282C-106	Sequence 106, App
	41	12.6	57.3	61	5	PCT-US92-06821A-52	Sequence 52, Appl
	42	12.6	57.3	76	1	US-08-442-572-37	Sequence 37, Appl
	43	12.6	57.3	76	1	US-08-361-795-37	Sequence 37, Appl
	44	12.6	57.3	76	5	PCT-US95-05600-120	Sequence 120, App
	45	12.6	57.3	77	1	US-08-447-169A-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1
US-09-092-314-2
; Sequence 2, Application US/09092314
; Patent No. 6225292
; GENERAL INFORMATION:
; APPLICANT: Raz, Eyal
; APPLICANT: Roman, Mark
; TITLE OF INVENTION: Inhibitors of DNA Immunostimulatory
; TITLE OF INVENTION: Sequence Activity
; Patent No. 6225292
; FILE REFERENCE: 6510-173US1
; CURRENT APPLICATION NUMBER: US/09/092,314
; CURRENT FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/048,794
; PRIOR FILING DATE: 1997-06-06
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Oligonucleotide
US-09-092-314-2

Query Match 92.7%; Score 20.4; DB 4; Length 22;
Best Local Similarity 95.5%; Pred. No. 0.049;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TGACTGTGAACGTTCGAGATGA 22
||||||||||||||| |||||||
Db 1 tgactgtgaacgttagagatga 22

RESULT 2
US-09-092-314-1
; Sequence 1, Application US/09092314
; Patent No. 6225292
; GENERAL INFORMATION:
; APPLICANT: Raz, Eyal
; APPLICANT: Roman, Mark
; TITLE OF INVENTION: Inhibitors of DNA Immunostimulatory
; TITLE OF INVENTION: Sequence Activity
; Patent No. 6225292
; FILE REFERENCE: 6510-173US1
; CURRENT APPLICATION NUMBER: US/09/092,314
; CURRENT FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/048,794
; PRIOR FILING DATE: 1997-06-06

```
;  NUMBER OF SEQ ID NOS:  11
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 22
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Oligonucleotide
US-09-092-314-1


  Query Match             85.5%;  Score 18.8;  DB 4;  Length 22;
  Best Local Similarity   90.9%;  Pred. No. 0.33;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          |||||||||| ||| |||||||
Db      1 tgactgtgaaggttagagatga 22


RESULT   3
US-09-092-314-3
; Sequence 3, Application US/09092314
; Patent No. 6225292
; GENERAL INFORMATION:
;  APPLICANT:  Raz, Eyal
;  APPLICANT:  Roman, Mark
;  TITLE OF INVENTION:  Inhibitors of DNA Immunostimulatory
;  TITLE OF INVENTION:  Sequence Activity
; Patent No. 6225292
;  FILE REFERENCE:  6510-173US1
;  CURRENT APPLICATION NUMBER: US/09/092,314
;  CURRENT FILING DATE:  1998-06-05
;  PRIOR APPLICATION NUMBER:  60/048,794
;  PRIOR FILING DATE:  1997-06-06
;  NUMBER OF SEQ ID NOS:  11
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 22
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Oligonucleotide
US-09-092-314-3


  Query Match             85.5%;  Score 18.8;  DB 4;  Length 22;
  Best Local Similarity   90.9%;  Pred. No. 0.33;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          |||||||||| || |||||||
Db      1 tgactgtgaaccttagagatga 22


RESULT   4
US-09-092-314-10
; Sequence 10, Application US/09092314
; Patent No. 6225292
; GENERAL INFORMATION:
;  APPLICANT:  Raz, Eyal
;  APPLICANT:  Roman, Mark
;  TITLE OF INVENTION:  Inhibitors of DNA Immunostimulatory
;  TITLE OF INVENTION:  Sequence Activity
; Patent No. 6225292
;  FILE REFERENCE:  6510-173US1
;  CURRENT APPLICATION NUMBER: US/09/092,314
;  CURRENT FILING DATE:  1998-06-05
;  PRIOR APPLICATION NUMBER:  60/048,794
;  PRIOR FILING DATE:  1997-06-06
;  NUMBER OF SEQ ID NOS:  11
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 10
;   LENGTH: 22
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Oligonucleotide
US-09-092-314-10


  Query Match             85.5%;  Score 18.8;  DB 4;  Length 22;
  Best Local Similarity   90.9%;  Pred. No. 0.33;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          |||||||||| ||| |||||||
Db      1 tgactgtgaatgttagagatga 22


RESULT   5
US-09-092-314-4
; Sequence 4, Application US/09092314
; Patent No. 6225292
; GENERAL INFORMATION:
;  APPLICANT:  Raz, Eyal
;  APPLICANT:  Roman, Mark
;  TITLE OF INVENTION:  Inhibitors of DNA Immunostimulatory
;  TITLE OF INVENTION:  Sequence Activity
; Patent No. 6225292
;  FILE REFERENCE:  6510-173US1
;  CURRENT APPLICATION NUMBER: US/09/092,314
;  CURRENT FILING DATE:  1998-06-05
;  PRIOR APPLICATION NUMBER:  60/048,794
;  PRIOR FILING DATE:  1997-06-06
;  NUMBER OF SEQ ID NOS:  11
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 22
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Oligonucleotide
US-09-092-314-4


  Query Match             78.2%;  Score 17.2;  DB 4;  Length 22;
  Best Local Similarity   86.4%;  Pred. No. 2.2;
  Matches   19;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          |||||||||| || |||||||
Db      1 tgactgtgaagcttagagatga 22


RESULT   6
US-09-092-314-5
; Sequence 5, Application US/09092314
; Patent No. 6225292
; GENERAL INFORMATION:
;  APPLICANT:  Raz, Eyal
;  APPLICANT:  Roman, Mark
;  TITLE OF INVENTION:  Inhibitors of DNA Immunostimulatory
;  TITLE OF INVENTION:  Sequence Activity
; Patent No. 6225292
;  FILE REFERENCE:  6510-173US1
;  CURRENT APPLICATION NUMBER: US/09/092,314
;  CURRENT FILING DATE:  1998-06-05
;  PRIOR APPLICATION NUMBER:  60/048,794
;  PRIOR FILING DATE:  1997-06-06
;  NUMBER OF SEQ ID NOS:  11
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 22
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
```

```
;   FEATURE:
;   OTHER INFORMATION: Oligonucleotide
US-09-092-314-5


  Query Match             70.9%;  Score 15.6;  DB 4;  Length 22;
  Best Local Similarity   81.8%;  Pred. No. 15;
  Matches   18;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           ||||||||  | || |||||||
Db       1 tgactgtgttccttagagatga 22


RESULT   7
US-09-092-314-7
; Sequence 7, Application US/09092314
; Patent No. 6225292
; GENERAL INFORMATION:
;  APPLICANT:  Raz, Eyal
;  APPLICANT:  Roman, Mark
;  TITLE OF INVENTION:  Inhibitors of DNA Immunostimulatory
;  TITLE OF INVENTION:  Sequence Activity
; Patent No. 6225292
;  FILE REFERENCE:  6510-173US1
;  CURRENT APPLICATION NUMBER: US/09/092,314
;  CURRENT FILING DATE:  1998-06-05
;  PRIOR APPLICATION NUMBER:  60/048,794
;  PRIOR FILING DATE:  1997-06-06
;  NUMBER OF SEQ ID NOS:  11
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 7
;   LENGTH: 22
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Oligonucleotide
US-09-092-314-7


  Query Match             70.9%;  Score 15.6;  DB 4;  Length 22;
  Best Local Similarity   81.8%;  Pred. No. 15;
  Matches   18;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           ||||||||  | ||  |||||||
Db       1 tgactgtggaggtcagagatga 22


RESULT   8
US-09-092-314-8
; Sequence 8, Application US/09092314
; Patent No. 6225292
; GENERAL INFORMATION:
;  APPLICANT:  Raz, Eyal
;  APPLICANT:  Roman, Mark
;  TITLE OF INVENTION:  Inhibitors of DNA Immunostimulatory
;  TITLE OF INVENTION:  Sequence Activity
; Patent No. 6225292
;  FILE REFERENCE:  6510-173US1
;  CURRENT APPLICATION NUMBER: US/09/092,314
;  CURRENT FILING DATE:  1998-06-05
;  PRIOR APPLICATION NUMBER:  60/048,794
;  PRIOR FILING DATE:  1997-06-06
;  NUMBER OF SEQ ID NOS:  11
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 8
;   LENGTH: 22
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Oligonucleotide
US-09-092-314-8


  Query Match             70.9%;  Score 15.6;  DB 4;  Length 22;
  Best Local Similarity   81.8%;  Pred. No. 15;
  Matches   18;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           ||||||||  | | |||||||
Db       1 tgactgtgagggctagagatga 22


RESULT   9
US-08-399-412A-58
; Sequence 58, Application US/08399412A
; Patent No. 5622828
;  GENERAL INFORMATION:
;    APPLICANT:  Parma, David
;    APPLICANT:  Gold, Larry
;    TITLE OF INVENTION:  High-Affinity Oligonucleotide
;    TITLE OF INVENTION:  Ligands To Secretory Phospholipase
;    TITLE OF INVENTION:  A2 (sPLA2)
;    NUMBER OF SEQUENCES:  122
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Swanson & Bratschun, L.L.C.
;      STREET:  8400 E. Prentice Avenue, Suite 200
;      CITY:  Englewood
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 1.44 MB storage
;      COMPUTER:  IBM compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/399,412A
;      FILING DATE:  6-MARCH-1995
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/714,131
;      FILING DATE:  10-JUNE-1991
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/536,428
;      FILING DATE:  11-JUNE-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/964,624
;      FILING DATE:  21-OCTOBER-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Julie L. Bernard
;      REGISTRATION NUMBER:  36,450
;      REFERENCE/DOCKET NUMBER:  NEX27
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 793-3333
;      TELEFAX:  (303) 793-3433
;  INFORMATION FOR SEQ ID NO:  58:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  77 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-399-412A-58


  Query Match             63.6%;  Score 14;  DB 1;  Length 77;
  Best Local Similarity   59.1%;  Pred. No. 1.2e+02;
  Matches   13;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           :| |   |||||::||| |:||
Db      42 UGCCCACGAACGUUCGACAUGA 63
```

```
RESULT 10
US-08-952-793-258
; Sequence 258, Application US/08952793
; Patent No. 6280932
;  GENERAL INFORMATION:
;    APPLICANT:  PARMA, et al.
;    TITLE OF INVENTION:  HIGH AFFINITY NUCLEIC ACID LIGANDS
;    TITLE OF INVENTION:  TO LECTINS
;    NUMBER OF SEQUENCES:  390
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Swanson & Bratschun, L.L.C.
;      STREET:  8400 E. Prentice Avenue, Suite 200
;      CITY:  Englewood
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3 1/2 diskette, 1.44 MB
;      COMPUTER:  IBM pc compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 6.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/952,793
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US96/09455
;      FILING DATE:  05-JUNE-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/479,724
;      FILING DATE:  07-JUNE-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/472,256
;      FILING DATE:  07-JUNE-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/472,255
;      FILING DATE:  07-JUNE-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/477,829
;      FILING DATE:  07-JUNE-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Barry J. Swanson
;      REGISTRATION NUMBER:  33,215
;      REFERENCE/DOCKET NUMBER:  NEX40C/PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 793-3333
;      TELEFAX:  (303) 793-3433
;  INFORMATION FOR SEQ ID NO:  258:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  95 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  RNA
;    FEATURE:
;      OTHER INFORMATION:  All C's are 2'-NH2 cytosine
;    FEATURE:
;      OTHER INFORMATION:  All U's are 2'-NH2 uracil
US-08-952-793-258


  Query Match           63.6%;  Score 14;  DB 4;  Length 95;
  Best Local Similarity  59.1%;  Pred. No. 1.2e+02;
  Matches   13;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           :|||:  ||| |::||| | ||
Db      60 UGACUCCGAAAGUUCGACAGGA 81


RESULT 11
PCT-US96-09455A-258
; Sequence 258, Application PC/TUS9609455A
;  GENERAL INFORMATION:
;    APPLICANT:  PARMA, et al.
;    TITLE OF INVENTION:  HIGH AFFINITY NUCLEIC ACID
;    TITLE OF INVENTION:  LIGANDS TO LECTINS
;    NUMBER OF SEQUENCES:  390
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Swanson & Bratschun, L.L.C.
;      STREET:  8400 E. Prentice Avenue, Suite 200
;      CITY:  Englewood
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3 1/2 diskette, 1.44 MB
;      COMPUTER:  IBM pc compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 6.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US96/09455A
;      FILING DATE:  05 JUNE 1996
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/479,724
;      FILING DATE:  07-JUNE-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/472,256
;      FILING DATE:  07-JUNE-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/472,255
;      FILING DATE:  07-JUNE-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/477,829
;      FILING DATE:  07-JUNE-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Barry J. Swanson
;      REGISTRATION NUMBER:  33,215
;      REFERENCE/DOCKET NUMBER:  NEX40C/PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 793-3333
;      TELEFAX:  (303) 793-3433
;  INFORMATION FOR SEQ ID NO:  258:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  95 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  RNA
;    FEATURE:
;      OTHER INFORMATION:  All C's are 2'-NH2 cytosine
;    FEATURE:
;      OTHER INFORMATION:  All U's are 2'-NH2 uracil
PCT-US96-09455A-258


  Query Match           63.6%;  Score 14;  DB 5;  Length 95;
  Best Local Similarity  59.1%;  Pred. No. 1.2e+02;
  Matches   13;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           :|||:  ||| |::||| | ||
Db      60 UGACUCCGAAAGUUCGACAGGA 81


RESULT 12
US-08-384-708A-195
; Sequence 195, Application US/08384708A
; Patent No. 5639868
;  GENERAL INFORMATION:
;    APPLICANT:  Gold, Larry
;    APPLICANT:  Janjic, Nebojsa
;    TITLE OF INVENTION:  High-Affinity RNA Ligands of Basic
;    TITLE OF INVENTION:  Fibroblast Growth Factors
;    NUMBER OF SEQUENCES:  227
```

```
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Swanson & Bratschun, L.L.C.
;      STREET:  8400 E. Prentice Avenue, Suite 200
;      CITY:  Englewood
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 1.44 MG storage
;      COMPUTER:  IBM compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/384,708A
;      FILING DATE:  02-FEBRUARY-1995
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/195,005
;      FILING DATE:  10-FEBRUARY-1994
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/714,131
;      FILING DATE:  10-JUNE-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/536,428
;      FILING DATE:  11-JUNE-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Barry J. Swanson
;      REGISTRATION NUMBER:  33,215
;      REFERENCE/DOCKET NUMBER:  NEX07/D
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 793-3333
;      TELEFAX:  (303) 793-3433
;  INFORMATION FOR SEQ ID NO:  195:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  77 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-384-708A-195


  Query Match             61.8%;  Score 13.6;  DB 1;  Length 77;
  Best Local Similarity   55.0%;  Pred. No. 1.9e+02;
  Matches   11;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       3 ACTGTGAACGTTCGAGATGA 22
           ||:|:|  | ::||| |:||
Db      44 ACUGUGCCCCUUCGACAUGA 63


RESULT  13
US-08-687-421-287
; Sequence 287, Application US/08687421
; Patent No. 6177557
;  GENERAL INFORMATION:
;    APPLICANT:  Gold, Larry
;    APPLICANT:  Janjic, Nebojsa
;    APPLICANT:  Tasset, Diane
;    TITLE OF INVENTION:  HIGH-AFFINITY LIGANDS OF BASIC
;    TITLE OF INVENTION:  FIBROBLAST GROWTH FACTOR AND
;    TITLE OF INVENTION:  THROMBIN
;    NUMBER OF SEQUENCES:  445
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Swanson & Bratschun, L.L.C.
;      STREET:  8400 E. Prentice Avenue, Suite 200
;      CITY:  Englewood
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 1.44 MB storage
;      COMPUTER:  IBM compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 6.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/687,421
;      FILING DATE:  08-MAY-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/195,005
;      FILING DATE:  10-FEBRUARY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:  22-APRIL-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/219,012
;      FILING DATE:  28-MARCH-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/973,333
;      FILING DATE:  11-NOVEMBER-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/714,131
;      FILING DATE:  10-JUNE-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/536,428
;      FILING DATE:  11-JUNE-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Barry J. Swanson
;      REGISTRATION NUMBER:  33,215
;      REFERENCE/DOCKET NUMBER:  NEX07/PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 793-3333
;      TELEFAX:  (303) 793-3433
;  INFORMATION FOR SEQ ID NO:  287:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  77 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-687-421-287


  Query Match             61.8%;  Score 13.6;  DB 4;  Length 77;
  Best Local Similarity   55.0%;  Pred. No. 1.9e+02;
  Matches   11;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       3 ACTGTGAACGTTCGAGATGA 22
           ||:|:|  | ::||| |:||
Db      44 ACUGUGCCCCUUCGACAUGA 63


RESULT  14
US-08-210-222-11/c
; Sequence 11, Application US/08210222
; Patent No. 5599917
;  GENERAL INFORMATION:
;    APPLICANT:  Coppola, George R.
;    APPLICANT:  Beutel, Bruce A.
;    APPLICANT:  Bertelsen, Arthur H.
;    TITLE OF INVENTION:  Inhibition of Interferon-  with Oligonucleotides
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carella, Byrne, Bain, Gilfillan,
;      ADDRESSEE:  Cecchi, Stewart & Olstein
;      STREET:  6 Becker Farm Road
;      CITY:  Roseland
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch diskette
;      COMPUTER:  IBM
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
```

```
;       APPLICATION NUMBER:  US/08/210,222
;       FILING DATE:  Unassigned
;       CLASSIFICATION:  514
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Herron, Charles J.
;       REGISTRATION NUMBER:  28,019
;       REFERENCE/DOCKET NUMBER:  23550-114
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  201-994-1700
;       TELEFAX:  201-994-1744
;   INFORMATION FOR SEQ ID NO:  11:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  97 BASES
;       TYPE:  NUCLEIC ACID
;       STRANDEDNESS:  SINGLE
;       TOPOLOGY:  LINEAR
;     HYPOTHETICAL:  NO
US-08-210-222-11


  Query Match             61.8%;  Score 13.6;  DB 1;  Length 97;
  Best Local Similarity   80.0%;  Pred. No. 2e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       3 ACTGTGAACGTTCGAGATGA 22
           ||||||| |  |||||| ||
Db      97 ACTGTGACCTCTCGAGACGA 78


RESULT  15
US-08-210-222-8/c
; Sequence 8, Application US/08210222
; Patent No. 5599917
;  GENERAL INFORMATION:
;    APPLICANT:  Coppola, George R.
;    APPLICANT:  Beutel, Bruce A.
;    APPLICANT:  Bertelsen, Arthur H.
;    TITLE OF INVENTION:  Inhibition of Interferon-  with Oligonucleotides
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carella, Byrne, Bain, Gilfillan,
;      ADDRESSEE:  Cecchi, Stewart & Olstein
;      STREET:  6 Becker Farm Road
;      CITY:  Roseland
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch diskette
;      COMPUTER:  IBM
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/210,222
;      FILING DATE:  Unassigned
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Herron, Charles J.
;      REGISTRATION NUMBER:  28,019
;      REFERENCE/DOCKET NUMBER:  23550-114
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  98 BASES
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  SINGLE
;      TOPOLOGY:  LINEAR
;    HYPOTHETICAL:  NO
US-08-210-222-8


  Query Match             61.8%;  Score 13.6;  DB 1;  Length 98;
  Best Local Similarity   80.0%;  Pred. No. 2e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       3 ACTGTGAACGTTCGAGATGA 22
           ||||||| |  |||||| |||
Db      98 ACTGTGACCTCTCGAGGTGA 79


RESULT  16
US-08-210-222-19/c
; Sequence 19, Application US/08210222
; Patent No. 5599917
;  GENERAL INFORMATION:
;    APPLICANT:  Coppola, George R.
;    APPLICANT:  Beutel, Bruce A.
;    APPLICANT:  Bertelsen, Arthur H.
;    TITLE OF INVENTION:  Inhibition of Interferon-  with Oligonucleotides
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carella, Byrne, Bain, Gilfillan,
;      ADDRESSEE:  Cecchi, Stewart & Olstein
;      STREET:  6 Becker Farm Road
;      CITY:  Roseland
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch diskette
;      COMPUTER:  IBM
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/210,222
;      FILING DATE:  Unassigned
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Herron, Charles J.
;      REGISTRATION NUMBER:  28,019
;      REFERENCE/DOCKET NUMBER:  23550-114
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  98 BASES
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  SINGLE
;      TOPOLOGY:  LINEAR
;    HYPOTHETICAL:  NO
US-08-210-222-19


  Query Match             61.8%;  Score 13.6;  DB 1;  Length 98;
  Best Local Similarity   80.0%;  Pred. No. 2e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       3 ACTGTGAACGTTCGAGATGA 22
           ||||||| |  |||||| ||
Db      98 ACTGTGACCTCTCGAGACGA 79


RESULT  17
US-08-210-222-22/c
; Sequence 22, Application US/08210222
; Patent No. 5599917
;  GENERAL INFORMATION:
;    APPLICANT:  Coppola, George R.
;    APPLICANT:  Beutel, Bruce A.
;    APPLICANT:  Bertelsen, Arthur H.
;    TITLE OF INVENTION:  Inhibition of Interferon-  with Oligonucleotides
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
```

```
;      ADDRESSEE:  Carella, Byrne, Bain, Gilfillan,
;      ADDRESSEE:  Cecchi, Stewart & Olstein
;      STREET:  6 Becker Farm Road
;      CITY:  Roseland
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch diskette
;      COMPUTER:  IBM
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/210,222
;      FILING DATE:  Unassigned
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Herron, Charles J.
;      REGISTRATION NUMBER:  28,019
;      REFERENCE/DOCKET NUMBER:  23550-114
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  98 BASES
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  SINGLE
;      TOPOLOGY:  LINEAR
;    HYPOTHETICAL:  NO
US-08-210-222-22


  Query Match             61.8%;  Score 13.6;  DB 1;  Length 98;
  Best Local Similarity   80.0%;  Pred. No. 2e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       3 ACTGTGAACGTTCGAGATGA 22
           ||||||| |  ||||| |||
Db      98 ACTGTGACCTCTCGAGGTGA 79


RESULT  18
US-08-210-222-24/c
; Sequence 24, Application US/08210222
; Patent No. 5599917
;  GENERAL INFORMATION:
;    APPLICANT:  Coppola, George R.
;    APPLICANT:  Beutel, Bruce A.
;    APPLICANT:  Bertelsen, Arthur H.
;    TITLE OF INVENTION:  Inhibition of Interferon-  with Oligonucleotides
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carella, Byrne, Bain, Gilfillan,
;      ADDRESSEE:  Cecchi, Stewart & Olstein
;      STREET:  6 Becker Farm Road
;      CITY:  Roseland
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch diskette
;      COMPUTER:  IBM
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/210,222
;      FILING DATE:  Unassigned
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Herron, Charles J.
;      REGISTRATION NUMBER:  28,019
;      REFERENCE/DOCKET NUMBER:  23550-114
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  98 BASES
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  SINGLE
;      TOPOLOGY:  LINEAR
;    HYPOTHETICAL:  NO
US-08-210-222-24


  Query Match             61.8%;  Score 13.6;  DB 1;  Length 98;
  Best Local Similarity   80.0%;  Pred. No. 2e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       3 ACTGTGAACGTTCGAGATGA 22
           ||||||| |  ||||| ||
Db      98 ACTGTGACCTCTCGAGAGGA 79


RESULT  19
US-08-633-768A-12/c
; Sequence 12, Application US/08633768A
; Patent No. 6013504
;  GENERAL INFORMATION:
;    APPLICANT:  YU, SHUKUN
;    APPLICANT:  BOJSEN, KIRSTEN
;    APPLICANT:  KRAGH, KARSTEN
;    APPLICANT:  BOJKO, MAJA
;    APPLICANT:  NIELSEN, JOHN
;    APPLICANT:  MARCUSSEN, JAN
;    TITLE OF INVENTION:  ALPHA-1,4-GLUCAN LYASE FROM
;    TITLE OF INVENTION:  A FUNGUS INFECTED ALGAE, ITS PURIFICATION
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Knobbe, Martens, Olson & Bear
;      STREET:  620 Newport Center Drive  16th Floor
;      CITY:  Newport Beach
;      STATE:  CA
;      COUNTRY:  U.S.A.
;      ZIP:  92660
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/633,768A
;      FILING DATE:  02-JUL-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  9321301.5
;      FILING DATE:  15-OCT-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Altman, Daniel E
;      REGISTRATION NUMBER:  34,115
;      REFERENCE/DOCKET NUMBER:  DYOU7.001APC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  714-760-0404
;      TELEFAX:  714-760-9502
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  71 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
US-08-633-768A-12
```

```
  Query Match             60.9%;  Score 13.4;  DB 3;  Length 71;
  Best Local Similarity   93.3%;  Pred. No. 2.4e+02;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 CTGTGAACGTTCGAG 18
           ||| ||||||||||||
Db      29 CTGCGAACGTTCGAG 15


RESULT  20
US-09-386-607-6
; Sequence 6, Application US/09386607
; Patent No. 6162628
; GENERAL INFORMATION:
;  APPLICANT: Cherry, Joel
;  APPLICANT: Svendsen, Allan
;  APPLICANT: Andersen, Carsten
;  APPLICANT: Beier, Lars
;  APPLICANT: Frandsen, Torben
;  TITLE OF INVENTION: Maltogenic Alpha-Amylase Variants
;  FILE REFERENCE: 5443.414-US
;  CURRENT APPLICATION NUMBER: US/09/386,607
;  CURRENT FILING DATE: 1999-08-31
;  EARLIER APPLICATION NUMBER: DK98/00269
;  EARLIER FILING DATE: 1998-02-27
;  EARLIER APPLICATION NUMBER: 60/077,795
;  EARLIER FILING DATE: 1998-03-12
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
;   LENGTH: 36
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: F 284D Primer
US-09-386-607-6


  Query Match             60.0%;  Score 13.2;  DB 4;  Length 36;
  Best Local Similarity   83.3%;  Pred. No. 2.8e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       5 TGTGAACGTTCGAGATGA 22
           ||| || |||| ||||||
Db       3 tgtcaatgttctagatga 20


RESULT  21
US-08-400-440A-19
; Sequence 19, Application US/08400440A
; Patent No. 5705337
;  GENERAL INFORMATION:
;    APPLICANT:  GOLD et al.
;    TITLE OF INVENTION:  SYSTEMATIC EVOLUTION OF LIGANDS BY
;    TITLE OF INVENTION:  EXPONENTIAL ENRICHMENT:  CHEMI-
;    TITLE OF INVENTION:  SELEX
;    NUMBER OF SEQUENCES:  104
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Swanson & Bratschun, L.L.C.
;      STREET:  8400 E. Prentice Avenue, Suite 200
;      CITY:  Englewood
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3 1/2 diskette, 1.44 MG
;      COMPUTER:  IBM pc compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/400,440A
;      FILING DATE:  08 MARCH 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/714,131
;      FILING DATE:  10-JUNE-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/536,428
;      FILING DATE:  11-JUNE-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/117,991
;      FILING DATE:  8-SEPTEMBER-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/123,935
;      FILING DATE:  17-SEPTEMBER-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/199,507
;      FILING DATE:  22-FEBRUARY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/234,997
;      FILING DATE:  28-APRIL-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/309,245
;      FILING DATE:  20-SEPTEMBER-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Barry J. Swanson
;      REGISTRATION NUMBER:  33,215
;      REFERENCE/DOCKET NUMBER:  NEX28
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 793-3333
;      TELEFAX:  (303) 793-3433
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  77 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-400-440A-19


  Query Match             60.0%;  Score 13.2;  DB 1;  Length 77;
  Best Local Similarity   61.1%;  Pred. No. 3.1e+02;
  Matches   11;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       5 TGTGAACGTTCGAGATGA 22
           :| | |||::||| |:||
Db      46 UGCGCACGUUCGACAUGA 63


RESULT  22
US-08-463-093A-19
; Sequence 19, Application US/08463093A
; Patent No. 5763595
;  GENERAL INFORMATION:
;    APPLICANT:  GOLD et al.
;    TITLE OF INVENTION:  SYSTEMATIC EVOLUTION OF LIGANDS BY
;    TITLE OF INVENTION:  EXPONENTIAL ENRICHMENT:  CHEMI-
;    TITLE OF INVENTION:  SELEX
;    NUMBER OF SEQUENCES:  104
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Swanson & Bratschun, L.L.C.
;      STREET:  8400 E. Prentice Avenue, Suite 200
;      CITY:  Englewood
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3 1/2 diskette, 1.44 MB
;      COMPUTER:  IBM pc compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 6.0 (a) For Windows
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/463,093A
;      FILING DATE:  05-JUNE-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  08/400,440
;      FILING DATE:  08-MARCH-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/714,131
;      FILING DATE:  10-JUNE-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/536,428
;      FILING DATE:  11-JUNE-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/117,991
;      FILING DATE:  8-SEPTEMBER-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/123,935
;      FILING DATE:  17-SEPTEMBER-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/199,507
;      FILING DATE:  22-FEBRUARY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/234,997
;      FILING DATE:  28-APRIL-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/309,245
;      FILING DATE:  20-SEPTEMBER-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Barry J. Swanson
;      REGISTRATION NUMBER:  33,215
;      REFERENCE/DOCKET NUMBER:  NEX28/C
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 793-3333
;      TELEFAX:  (303) 793-3433
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  77 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-463-093A-19


  Query Match             60.0%;  Score 13.2;  DB 1;  Length 77;
  Best Local Similarity   61.1%;  Pred. No. 3.1e+02;
  Matches   11;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       5 TGTGAACGTTCGAGATGA 22
           :| | |||::||| |:||
Db      46 UGCGCACGUUCGACAUGA 63


RESULT  23
US-08-460-888A-19
; Sequence 19, Application US/08460888A
; Patent No. 5962219
;  GENERAL INFORMATION:
;    APPLICANT:  GOLD et al.
;    TITLE OF INVENTION:  SYSTEMATIC EVOLUTION OF LIGANDS BY
;    TITLE OF INVENTION:  EXPONENTIAL ENRICHMENT:  CHEMI-
;    TITLE OF INVENTION:  SELEX
;    NUMBER OF SEQUENCES:  104
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Swanson & Bratschun, L.L.C.
;      STREET:  8400 E. Prentice Avenue, Suite 200
;      CITY:  Englewood
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3 1/2 diskette, 1.44 MB
;      COMPUTER:  IBM pc compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 6.0 (a) For Windows
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/460,888A
;      FILING DATE:  05-JUNE-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/400,440
;      FILING DATE:  08-MARCH-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/714,131
;      FILING DATE:  10-JUNE-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/536,428
;      FILING DATE:  11-JUNE-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/117,991
;      FILING DATE:  8-SEPTEMBER-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/123,935
;      FILING DATE:  17-SEPTEMBER-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/199,507
;      FILING DATE:  22-FEBRUARY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/234,997
;      FILING DATE:  28-APRIL-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/309,245
;      FILING DATE:  20-SEPTEMBER-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Barry J. Swanson
;      REGISTRATION NUMBER:  33,215
;      REFERENCE/DOCKET NUMBER:  NEX28/C
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 793-3333
;      TELEFAX:  (303) 793-3433
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  77 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-460-888A-19


  Query Match             60.0%;  Score 13.2;  DB 2;  Length 77;
  Best Local Similarity   61.1%;  Pred. No. 3.1e+02;
  Matches   11;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       5 TGTGAACGTTCGAGATGA 22
           :| | |||::||| |:||
Db      46 UGCGCACGUUCGACAUGA 63


RESULT  24
US-08-894-578-19
; Sequence 19, Application US/08894578
; Patent No. 5998142
;  GENERAL INFORMATION:
;    APPLICANT:  GOLD et al.
;    TITLE OF INVENTION:  SYSTEMATIC EVOLUTION OF LIGANDS
;    TITLE OF INVENTION:  BY EXPONENTIAL ENRICHMENT:
;    TITLE OF INVENTION:  CHEMI-SELEX
;    NUMBER OF SEQUENCES:  226
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Swanson & Bratschun, L.L.C.
;      STREET:  8400 E. Prentice Avenue, Suite 200
;      CITY:  Englewood
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3 1/2 diskette, 1.44 MG
;      COMPUTER:  IBM pc compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 6.0
;    CURRENT APPLICATION DATA:
```

```
;       APPLICATION NUMBER:  US/08/894,578
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US96/03097
;       FILING DATE:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/714,131
;       FILING DATE:  10-JUNE-1991
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/536,428
;       FILING DATE:  11-JUNE-1990
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/117,991
;       FILING DATE:  8-SEPTEMBER-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/123,935
;       FILING DATE:  17-SEPTEMBER-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/199,507
;       FILING DATE:  22-FEBRUARY-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/234,997
;       FILING DATE:  28-APRIL-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/309,245
;       FILING DATE:  20-SEPTEMBER-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Barry J. Swanson
;       REGISTRATION NUMBER:  33,215
;       REFERENCE/DOCKET NUMBER:  NEX28/PCT
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (303) 793-3333
;       TELEFAX:  (303) 793-3433
;   INFORMATION FOR SEQ ID NO:  19:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  77 nucleotides
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  RNA
US-08-894-578-19


  Query Match             60.0%;  Score 13.2;  DB 2;  Length 77;
  Best Local Similarity   61.1%;  Pred. No. 3.1e+02;
  Matches   11;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       5 TGTGAACGTTCGAGATGA 22
           :| | |||::||| |:||
Db      46 UGCGCACGUUCGACAUGA 63


RESULT  25
US-09-412-017-19
; Sequence 19, Application US/09412017
; Patent No. 6300074
;  GENERAL INFORMATION:
;    APPLICANT:  GOLD et al.
;    TITLE OF INVENTION:  SYSTEMATIC EVOLUTION O
;    TITLE OF INVENTION:  EXPONENTIAL ENRICHMENT:  CHEMI-
;    TITLE OF INVENTION:  SELEX
;    NUMBER OF SEQUENCES:  104
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Swanson & Bratschun, L.L.C.
;      STREET:  8400 E. Prentice Avenue, Suite 200
;      CITY:  Englewood
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3 1/2 diskette, 1.44 MB
;      COMPUTER:  IBM pc compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 8.0 For Windows
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/412,017
;      FILING DATE:  04-OCTOBER-1999
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/460,888
;      FILING DATE:  05-JUNE-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/400,440
;      FILING DATE:  08-MARCH-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/714,131
;      FILING DATE:  10-JUNE-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/536,428
;      FILING DATE:  11-JUNE-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/117,991
;      FILING DATE:  8-SEPTEMBER-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/123,935
;      FILING DATE:  17-SEPTEMBER-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/199,507
;      FILING DATE:  22-FEBRUARY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/234,997
;      FILING DATE:  28-APRIL-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/309,245
;      FILING DATE:  20-SEPTEMBER-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Barry J. Swanson
;      REGISTRATION NUMBER:  33,215
;      REFERENCE/DOCKET NUMBER:  NEX28/C3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 793-3333
;      TELEFAX:  (303) 793-3433
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  77 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-412-017-19


  Query Match             60.0%;  Score 13.2;  DB 4;  Length 77;
  Best Local Similarity   61.1%;  Pred. No. 3.1e+02;
  Matches   11;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       5 TGTGAACGTTCGAGATGA 22
           :| | |||::||| |:||
Db      46 UGCGCACGUUCGACAUGA 63


RESULT  26
US-08-210-222-4/c
; Sequence 4, Application US/08210222
; Patent No. 5599917
;  GENERAL INFORMATION:
;    APPLICANT:  Coppola, George R.
;    APPLICANT:  Beutel, Bruce A.
;    APPLICANT:  Bertelsen, Arthur H.
;    TITLE OF INVENTION:  Inhibition of Interferon-  with Oligonucleotides
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carella, Byrne, Bain, Gilfillan,
;      ADDRESSEE:  Cecchi, Stewart & Olstein
;      STREET:  6 Becker Farm Road
;      CITY:  Roseland
```

```
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch diskette
;      COMPUTER:  IBM
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/210,222
;      FILING DATE:  Unassigned
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Herron, Charles J.
;      REGISTRATION NUMBER:  28,019
;      REFERENCE/DOCKET NUMBER:  23550-114
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  97 BASES
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  SINGLE
;      TOPOLOGY:  LINEAR
;    HYPOTHETICAL:  NO
US-08-210-222-4


  Query Match             60.0%;  Score 13.2;  DB 1;  Length 97;
  Best Local Similarity   83.3%;  Pred. No. 3.2e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 ACTGTGAACGTTCGAGAT 20
           ||||||| |  |||||||
Db      97 ACTGTGACCTCTCGAGAT 80


RESULT  27
US-08-210-222-7/c
; Sequence 7, Application US/08210222
; Patent No. 5599917
;  GENERAL INFORMATION:
;    APPLICANT:  Coppola, George R.
;    APPLICANT:  Beutel, Bruce A.
;    APPLICANT:  Bertelsen, Arthur H.
;    TITLE OF INVENTION:  Inhibition of Interferon-  with Oligonucleotides
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carella, Byrne, Bain, Gilfillan,
;      ADDRESSEE:  Cecchi, Stewart & Olstein
;      STREET:  6 Becker Farm Road
;      CITY:  Roseland
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch diskette
;      COMPUTER:  IBM
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/210,222
;      FILING DATE:  Unassigned
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Herron, Charles J.
;      REGISTRATION NUMBER:  28,019
;      REFERENCE/DOCKET NUMBER:  23550-114
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  98 BASES
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  SINGLE
;      TOPOLOGY:  LINEAR
;    HYPOTHETICAL:  NO
US-08-210-222-7


  Query Match             60.0%;  Score 13.2;  DB 1;  Length 98;
  Best Local Similarity   83.3%;  Pred. No. 3.2e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 ACTGTGAACGTTCGAGAT 20
           ||||||| |  |||||||
Db      98 ACTGTGACCTCTCGAGAT 81


RESULT  28
US-09-017-612A-1/c
; Sequence 1, Application US/09017612A
; Patent No. 6194183
;  GENERAL INFORMATION:
;    APPLICANT:  Markvardsen, Peter
;    APPLICANT:  Bjornvad, Mads Eskelund
;    APPLICANT:  Mikkelsen, Frank
;    APPLICANT:  Diderichsen, Borge
;    TITLE OF INVENTION:  Phage Display For Detergent
;    TITLE OF INVENTION:  Enzyme Activity
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6194183o No. 6194183disk of No. 6194183th America, Inc.
;      STREET:  405 Lexington Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10174
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/017,612A
;      FILING DATE:  29-JAN-1998
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lambiris, Elias J
;      REGISTRATION NUMBER:  33,728
;      REFERENCE/DOCKET NUMBER:  4542.204-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-867-0123
;      TELEFAX:  212-878-9655
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  60 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-017-612A-1


  Query Match             59.1%;  Score 13;  DB 4;  Length 60;
  Best Local Similarity   76.2%;  Pred. No. 3.7e+02;
  Matches   16;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       2 GACTGTGAACGTTCGAGATGA 22
           | ||||| || |||| || ||
Db      29 GCCTGTGCACATTCGCGAGGA 9


RESULT  29
US-09-070-408-110
```

```
; Sequence 110, Application US/09070408
; Patent No. 6180341
; GENERAL INFORMATION:
;   APPLICANT: Iverson, Brent L.
;   APPLICANT: Georgiou, George
;   APPLICANT: Burks, Elizabeth A.
;   TITLE OF INVENTION: IN VITRO SCANNING SATURATION MUTAGENESIS
;   TITLE OF INVENTION: OF PROTEINS
;   NUMBER OF SEQUENCES: 132
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Arnold, White & Durkee
;     STREET: P.O. Box 4433
;     CITY: Houston
;     STATE: Texas
;     COUNTRY: USA
;     ZIP: 77210
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/09/070,408
;     FILING DATE: Concurrently Herewith
;     CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 60/045,409
;     FILING DATE: 01-MAY-1997
;   ATTORNEY/AGENT INFORMATION:
;     NAME: McMillian, Nabeela R.
;     REGISTRATION NUMBER: P-43,363
;     REFERENCE/DOCKET NUMBER: UTSB:593
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 512/418-3000
;     TELEFAX: 512/447-7577
; INFORMATION FOR SEQ ID NO: 110:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 31 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
US-09-070-408-110


  Query Match             58.2%;  Score 12.8;  DB 4;  Length 31;
  Best Local Similarity   87.5%;  Pred. No. 4.3e+02;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCG 16
           ||||  ||||||||||
Db      16 TGACCATGAACGTTCG 31


RESULT 30
US-08-403-762A-163
; Sequence 163, Application US/08403762A
; Patent No. 5703217
; GENERAL INFORMATION:
;   APPLICANT: MABILAT, Claude
;   APPLICANT: CHRISTEN, Richard
;   TITLE OF INVENTION: NUCLEOTIDE FRAGMENT OF THE 23S RIBOSOMAL
;   TITLE OF INVENTION: RNA OF MYCOBACTERIA, DERIVED PROBES AND PRIMERS, REAGENT
;   TITLE OF INVENTION: AND DETECTION METHOD
;   NUMBER OF SEQUENCES: 178
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: OLIFF & BERRIDGE
;     STREET: 700 South Washington Street, Suite 300
;     CITY: Alexandria
;     STATE: Virginia
;     COUNTRY: USA
;     ZIP: 22314
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/403,762A
;     FILING DATE: 23-MAR-1995
;     CLASSIFICATION: 536
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Berridge, William P
;     REGISTRATION NUMBER: 30,024
;     REFERENCE/DOCKET NUMBER: WPB 29658
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 703-836-6400
;     TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 163:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 36 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: rRNA
;   ORIGINAL SOURCE:
;     ORGANISM: M. PHLEI
;     STRAIN: A 247
;   POSITION IN GENOME:
;     MAP POSITION: 2126..2161, with respect to the numbering of
;     MAP POSITION: E. coli
US-08-403-762A-163


  Query Match             58.2%;  Score 12.8;  DB 1;  Length 36;
  Best Local Similarity   62.5%;  Pred. No. 4.4e+02;
  Matches   10;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 GACTGTGAACGTTCGA 17
             |||:|:|||  ::|||
Db       2 GACUGUGAAGCUUCGA 17


RESULT 31
US-08-447-169A-17
; Sequence 17, Application US/08447169A
; Patent No. 5811533
; GENERAL INFORMATION:
;   APPLICANT: JANJIC, N. and GOLD, L.
;   TITLE OF INVENTION: HIGH-AFFINITY OLIGONUCLEOTIDE
;   TITLE OF INVENTION: LIGANDS TO VASCULAR ENDOTHELIAL
;   TITLE OF INVENTION: GROWTH FACTOR (VEGF)
;   NUMBER OF SEQUENCES: 242
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Swanson & Bratschun, L.L.C.
;     STREET: 8400 E. Prentice Place, Suite 200
;     CITY: Englewood
;     STATE: Colorado
;     COUNTRY: USA
;     ZIP: 80111
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MG storage
;     COMPUTER: IBM compatible
;     OPERATING SYSTEM: MS-DOS
;     SOFTWARE: WordPerfect 5.1
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/447,169A
;     FILING DATE: 19-MAY-1995
;     CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 08/233,012
;     FILING DATE: 25-APRIL-1994
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 08/205,515
;     FILING DATE: 03-MARCH-1994
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 07/964,624
```

```
;      FILING DATE:  21-OCTOBER-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/714,131
;      FILING DATE:  10-JUNE-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/536,428
;      FILING DATE:  11-JUNE-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Barry J. Swanson
;      REGISTRATION NUMBER:  33,215
;      REFERENCE/DOCKET NUMBER:  NEX14
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 793-3333
;      TELEFAX:  (303) 793-3433
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  77 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-447-169A-17


  Query Match           58.2%;  Score 12.8;  DB 1;  Length 77;
  Best Local Similarity  62.5%;  Pred. No. 4.9e+02;
  Matches   10;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       7 TGAACGTTCGAGATGA 22
           :|| ||::||| |:||
Db      48 UGACCGUUCGACAUGA 63


RESULT  32
US-08-233-012C-17
; Sequence 17, Application US/08233012C
; Patent No. 5849479
;  GENERAL INFORMATION:
;    APPLICANT:  JANJIC, N. and GOLD, L.
;    TITLE OF INVENTION:  HIGH-AFFINITY OLIGONUCLEOTIDE
;    TITLE OF INVENTION:  LIGANDS TO VASCULAR
;    TITLE OF INVENTION:  ENDOTHELIAL GROWTH FACTOR
;    TITLE OF INVENTION:
;    NUMBER OF SEQUENCES:  146
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Swanson & Bratschun, L.L.C.
;      STREET:  8400 E. Prentice Place, Suite 200
;      CITY:  Englewood
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.44 MG
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/233,012C
;      FILING DATE:  25-APRIL-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/714,131
;      FILING DATE:  10-JUNE-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/536,428
;      FILING DATE:  11-JUNE-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/964,624
;      FILING DATE:  21-OCTOBER-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Barry J. Swanson
;      REGISTRATION NUMBER:  33,215
;      REFERENCE/DOCKET NUMBER:  NEX14
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 793-3333
;      TELEFAX:  (303) 793-3433
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  77 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-233-012C-17


  Query Match           58.2%;  Score 12.8;  DB 2;  Length 77;
  Best Local Similarity  62.5%;  Pred. No. 4.9e+02;
  Matches   10;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       7 TGAACGTTCGAGATGA 22
           :|| ||::||| |:||
Db      48 UGACCGUUCGACAUGA 63


RESULT  33
US-08-403-762A-149
; Sequence 149, Application US/08403762A
; Patent No. 5703217
;  GENERAL INFORMATION:
;    APPLICANT:  MABILAT, Claude
;    APPLICANT:  CHRISTEN, Richard
;    TITLE OF INVENTION:  NUCLEOTIDE FRAGMENT OF THE 23S RIBOSOMAL
;    TITLE OF INVENTION:  RNA OF MYCOBACTERIA, DERIVED PROBES AND PRIMERS, REAGENT
;    TITLE OF INVENTION:  AND DETECTION METHOD
;    NUMBER OF SEQUENCES:  178
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OLIFF & BERRIDGE
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  USA
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/403,762A
;      FILING DATE:  23-MAR-1995
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berridge, William P
;      REGISTRATION NUMBER:  30,024
;      REFERENCE/DOCKET NUMBER:  WPB 29658
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6400
;      TELEFAX:  703-836-2787
;  INFORMATION FOR SEQ ID NO:  149:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  26 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  rRNA
;    ORIGINAL SOURCE:
;      ORGANISM:  M. FORTUITUM
;      STRAIN:  CIP 140 410 001
;    POSITION IN GENOME:
;      MAP POSITION:  1854..1876, with respect to the numbering of
;      MAP POSITION:  E. coli
US-08-403-762A-149


  Query Match           57.3%;  Score 12.6;  DB 1;  Length 26;
  Best Local Similarity  52.6%;  Pred. No. 5.4e+02;
```

```
  Matches   10;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       4 CTGTGAACGTTCGAGATGA 22
           |:|: ||| ':::|| | :||
Db       8 CUGUUAACCUUCGGGGUGA 26


RESULT  34
US-08-308-952-18
; Sequence 18, Application US/08308952
; Patent No. 5837812
;  GENERAL INFORMATION:
;    APPLICANT:  Harrison, Leonard
;    APPLICANT:  Honeyman, Margot
;    APPLICANT:  Cram, David
;    APPLICANT:  Deaizpurua, Henry
;    TITLE OF INVENTION:  A METHOD FOR THE DIAGNOSIS AND TREATMENT
;    TITLE OF INVENTION:  OF GLUTAMIC ACID DECARBOXYLASE AUTOANTIGEN
;    TITLE OF INVENTION:  ASSOCIATED DISEASES
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Scully, Scott, Murphy & Presser
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/308,952
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  839,805
;      FILING DATE:  21-FEb-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S.
;      REGISTRATION NUMBER:  31,346
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;      TELEX:  230 901 SANS UR
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  27 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-308-952-18


  Query Match             57.3%;  Score 12.6;  DB 2;  Length 27;
  Best Local Similarity   78.9%;  Pred. No. 5.4e+02;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       4 CTGTGAACGTTCGAGATGA 22
           ||||||  |||| || |||
Db       8 CTGTGAGGGTTCCAGGTGA 26


RESULT  35
US-09-124-141-27
; Sequence 27, Application US/09124141
; Patent No. 6211352
; GENERAL INFORMATION:
;  APPLICANT: Harrison, Leonard
;  APPLICANT: Honeyman, Margot
;  APPLICANT: Cram, David
;  APPLICANT: De Aizpurua, Henry
;  TITLE OF INVENTION: A METHOD FOR THE DIAGNOSIS AND TREATMENT OF GLUTAMIC
;  TITLE OF INVENTION: ACID DECARBOXYLASE AUTOANTIGEN ASSOCIATED DISEASES
;  FILE REFERENCE: Phillips, Ormonde & Fitzpatrick
;  CURRENT APPLICATION NUMBER: US/09/124,141
;  CURRENT FILING DATE: 1998-07-29
;  EARLIER APPLICATION NUMBER: 08/308,952
;  EARLIER FILING DATE: 1994-09-20
;  EARLIER APPLICATION NUMBER: 07/839,805
;  EARLIER FILING DATE: 1992-02-21
;  NUMBER OF SEQ ID NOS: 34
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
;   LENGTH: 27
;   TYPE: DNA
;   ORGANISM: Unknown Organism
;   FEATURE:
;   OTHER INFORMATION: Description of Unknown Organism: Oligonucleotide
;   OTHER INFORMATION: Primer (RGAD4)
US-09-124-141-27


  Query Match             57.3%;  Score 12.6;  DB 4;  Length 27;
  Best Local Similarity   78.9%;  Pred. No. 5.4e+02;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       4 CTGTGAACGTTCGAGATGA 22
           ||||||  |||| || |||
Db       8 ctgtgagggttccaggtga 26


RESULT  36
US-08-153-799-12
; Sequence 12, Application US/08153799
; Patent No. 5766883
;  GENERAL INFORMATION:
;    APPLICANT:  Ballance, David J
;    APPLICANT:  Goodey, Andrew R
;    TITLE OF INVENTION:  Polypeptides
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  R Hain Swope, BOC Health Care Inc
;      STREET:  100 Mountain Avenue
;      CITY:  Murray Hill
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07974
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/153,799
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/847975
;      FILING DATE:  06-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 8909916.2
;      FILING DATE:  29-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB90/00650
;      FILING DATE:  26-APR-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/775952
;      FILING DATE:  29-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Swope, R Hain
;      REGISTRATION NUMBER:  24864
;      REFERENCE/DOCKET NUMBER:  92H832
;    TELECOMMUNICATION INFORMATION:
```

```
;       TELEPHONE:  (908) 665 2400
;       TELEFAX:  (908) 771 6159
;       TELEX:  219484
;   INFORMATION FOR SEQ ID NO:  12:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  36 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  6..11
;       OTHER INFORMATION:  /function= "KpnI site"
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  1..36
;       OTHER INFORMATION:  /product= "OLIGONUCLEOTIDE 9"
US-08-153-799-12


  Query Match             57.3%;  Score 12.6;  DB 1;  Length 36;
  Best Local Similarity   78.9%;  Pred. No. 5.6e+02;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       3 ACTGTGAACGTTCGAGATG 21
           |||||||  |||| | |||
Db      17 ACTGTGACTGTTCTAAATG 35


RESULT  37
US-09-017-612A-3/c
; Sequence 3, Application US/09017612A
; Patent No. 6194183
;  GENERAL INFORMATION:
;    APPLICANT:  Markvardsen, Peter
;    APPLICANT:  Bjornvad, Mads Eskelund
;    APPLICANT:  Mikkelsen, Frank
;    APPLICANT:  Diderichsen, Borge
;    TITLE OF INVENTION:  Phage Display For Detergent
;    TITLE OF INVENTION:  Enzyme Activity
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6194183o No. 6194183disk of No. 6194183th America, Inc.
;      STREET:  405 Lexington Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10174
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/017,612A
;      FILING DATE:  29-JAN-1998
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lambiris, Elias J
;      REGISTRATION NUMBER:  33,728
;      REFERENCE/DOCKET NUMBER:  4542.204-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-867-0123
;      TELEFAX:  212-878-9655
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  57 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-017-612A-3


  Query Match             57.3%;  Score 12.6;  DB 4;  Length 57;
  Best Local Similarity   78.9%;  Pred. No. 6e+02;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       4 CTGTGAACGTTCGAGATGA 22
           ||||| || |||| || ||
Db      27 CTGTGCACATTCGCGAGGA 9


RESULT  38
US-08-440-084-7/c
; Sequence 7, Application US/08440084
; Patent No. 5593835
;  GENERAL INFORMATION:
;    APPLICANT:  Rando, Robert R.
;    APPLICANT:  Wang, Yong
;    TITLE OF INVENTION:  METHODS AND KITS FOR RNA BINDING
;    TITLE OF INVENTION:  COMPOUNDS
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lappin & Kusmer
;      STREET:  200 State Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/440,084
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kerner, Ann-Louise
;      REGISTRATION NUMBER:  33,523
;      REFERENCE/DOCKET NUMBER:  HAZ-014
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-330-1300
;      TELEFAX:  617-330-1311
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  59 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  RNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-440-084-7


  Query Match             57.3%;  Score 12.6;  DB 1;  Length 59;
  Best Local Similarity   78.9%;  Pred. No. 6e+02;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGA 19
           | || |||||||| | |||
Db      25 TAACAGTGAACGTACAAGA 7


RESULT  39
PCT-US96-06669-7/c
; Sequence 7, Application PC/TUS9606669
;  GENERAL INFORMATION:
;    APPLICANT:  President and Fellows of Harvard College
;    TITLE OF INVENTION:  METHODS AND KITS FOR RNA BINDING
;    TITLE OF INVENTION:  COMPOUNDS
```

```
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lappin & Kusmer
;      STREET:  200 State Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US96/06669
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kerner, Ann-Louise
;      REGISTRATION NUMBER:  33,523
;      REFERENCE/DOCKET NUMBER:  HAZ-014PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-330-1300
;      TELEFAX:  617-330-1311
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  59 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  RNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
PCT-US96-06669-7


  Query Match             57.3%;  Score 12.6;  DB 5;  Length 59;
  Best Local Similarity   78.9%;  Pred. No. 6e+02;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGA 19
           | || |||||||| | |||
Db      25 TAACAGTGAACGTACAAGA 7


RESULT  40
US-07-744-282C-106
; Sequence 106, Application US/07744282C
; Patent No. 5521300
;  GENERAL INFORMATION:
;    APPLICANT:  Shah, Jyotsna S.
;    APPLICANT:  Nietupski, Raymond M.
;    APPLICANT:  Liu, Jing
;    TITLE OF INVENTION:  Oligonucleotides Complementary to
;    TITLE OF INVENTION:  Mycobacterial Nucleic Acids
;    NUMBER OF SEQUENCES:  127
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kevin M. Farrell, P.C.
;      STREET:  P.O. Box 999
;      CITY:  York Harbor
;      STATE:  ME
;      COUNTRY:  USA
;      ZIP:  03911
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/744,282C
;      FILING DATE:  August 13, 1991
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kevin M. Farrell
;      REGISTRATION NUMBER:  35,505
;      REFERENCE/DOCKET NUMBER:  GTR90-05
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (207) 363-0558
;      TELEFAX:  (207) 363-0528
;  INFORMATION FOR SEQ ID NO:  106:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  61 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  RNA (genomic)
US-07-744-282C-106


  Query Match             57.3%;  Score 12.6;  DB 1;  Length 61;
  Best Local Similarity   52.6%;  Pred. No. 6e+02;
  Matches   10;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       4 CTGTGAACGTTCGAGATGA 22
           |:|: ||| ::|| | :||
Db      12 CUGUUAACCUUCGGGGUGA 30


RESULT  41
PCT-US92-06821A-52
; Sequence 52, Application PC/TUS9206821A
;  GENERAL INFORMATION:
;    APPLICANT:  Shah, Jyotsna S.
;    APPLICANT:  Nietupski, Raymond M.
;    APPLICANT:  Liu, Jing
;    TITLE OF INVENTION:  Oligonucleotides Complementary to
;    TITLE OF INVENTION:  Mycobacterial Nucleic Acids
;    NUMBER OF SEQUENCES:  133
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amoco Corporation
;      STREET:  200 East Randolph Drive, P.O. Box 87703
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  U.S.A.
;      ZIP:  60680
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/06821A
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/744,282
;      FILING DATE:  13-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Galloway, Norval B.
;      REGISTRATION NUMBER:  33,595
;      REFERENCE/DOCKET NUMBER:  CN 5851
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-856-7180
;      TELEFAX:  312-856-4972
;  INFORMATION FOR SEQ ID NO:  52:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  61 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  RNA
PCT-US92-06821A-52


  Query Match             57.3%;  Score 12.6;  DB 5;  Length 61;
  Best Local Similarity   52.6%;  Pred. No. 6e+02;
  Matches   10;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;
```

```
Qy       4 CTGTGAACGTTCGAGATGA 22
           |:|: ||| ::|| | :||
Db      12 CUGUUAACCUUCGGGGUGA 30


RESULT  42
US-08-442-572-37
; Sequence 37, Application US/08442572
; Patent No. 5587468
;  GENERAL INFORMATION:
;    APPLICANT:  Allen, Patrick, and Gold, Larry
;    TITLE OF INVENTION:  High Affinity HIV
;    TITLE OF INVENTION:  Integrase Inhibitors
;    NUMBER OF SEQUENCES:  61
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Swanson and Bratschun, L.L.C.
;      STREET:  8400 East Prentice Avenue, Suite #200
;      CITY:  Denver
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 1.44 MG
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/442,572
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/361,795
;      FILING DATE:  21-DECEMBER-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/714,131
;      FILING DATE:  10-JUNE-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/536,428
;      FILING DATE:  11-JUNE-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/964,624
;      FILING DATE:  21-OCTOBER-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/177,991
;      FILING DATE:  8-SEPTEMBER-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/931,473
;      FILING DATE:  17-AUGUST-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Barry J. Swanson
;      REGISTRATION NUMBER:  33,215
;      REFERENCE/DOCKET NUMBER:  NEX25
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 793-3333
;      TELEFAX:  (303) 793-3433
;  INFORMATION FOR SEQ ID NO:  37:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  76
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-442-572-37


  Query Match             57.3%;  Score 12.6;  DB 1;  Length 76;
  Best Local Similarity   57.9%;  Pred. No. 6.2e+02;
  Matches   11;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       4 CTGTGAACGTTCGAGATGA 22
           |:  ||| |::||| |:||
Db      44 CUUAGAAAGUUCGACAUGA 62


RESULT  43
US-08-361-795-37
; Sequence 37, Application US/08361795
; Patent No. 5756287
;  GENERAL INFORMATION:
;    APPLICANT:  Allen, Patrick, and Gold, Larry
;    TITLE OF INVENTION:  High Affinity HIV
;    TITLE OF INVENTION:  Integrase Inhibitors
;    NUMBER OF SEQUENCES:  61
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Swanson and Bratschun, L.L.C.
;      STREET:  8400 East Prentice Avenue, Suite #200
;      CITY:  Denver
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 1.44 MG storage
;      COMPUTER:  IBM compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/361,795
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/714,131
;      FILING DATE:  10-JUNE-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/536,428
;      FILING DATE:  11-JUNE-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/964,624
;      FILING DATE:  21-OCTOBER-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/177,991
;      FILING DATE:  8-SEPTEMBER-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/931,473
;      FILING DATE:  17-AUGUST-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Barry J. Swanson
;      REGISTRATION NUMBER:  33,215
;      REFERENCE/DOCKET NUMBER:  NEX25
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 793-3333
;      TELEFAX:  (303) 793-3433
;  INFORMATION FOR SEQ ID NO:  37:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  76
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-361-795-37


  Query Match             57.3%;  Score 12.6;  DB 1;  Length 76;
  Best Local Similarity   57.9%;  Pred. No. 6.2e+02;
  Matches   11;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       4 CTGTGAACGTTCGAGATGA 22
           |:  ||| |::||| |:||
Db      44 CUUAGAAAGUUCGACAUGA 62


RESULT  44
PCT-US95-05600-120
; Sequence 120, Application PC/TUS9505600
;  GENERAL INFORMATION:
;    APPLICANT:  GOLD, LARRY
;    APPLICANT:  NIEUWLANDT, DAN
;    APPLICANT:  WECKER, MATTHEW
```

```
;    APPLICANT:  SCHNEIDER, DANIEL J.
;    APPLICANT:  FEIGON, JULI
;    APPLICANT:  ALLEN, PATRICK
;    APPLICANT:  SULLENGER, BRUCE A.
;    APPLICANT:  DOUDNA, JENNIFER, A.
;    TITLE OF INVENTION:  HIGH-AFFINITY LIGANDS OF
;    TITLE OF INVENTION:  INSULIIN RECEPTOR ANTIBODIES, TACHYKININ SUBSTANCE
;    TITLE OF INVENTION:  P, HIV INTEGRASE AND HIV-1 REVERSE TRANSCRIPTASE
;    NUMBER OF SEQUENCES:  239
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Swanson & Bratschun, L.L.C.
;      STREET:  8400 E. Prentice Avenue, Suite 200
;      CITY:  Englewood
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 1.44 MG
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/05600
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/238,863
;      FILING DATE:  06-MAY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/248,632
;      FILING DATE:  24-MAY-1994
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/303,362
;      FILING DATE:  09-SEPTEMBER-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/361,795
;      FILING DATE:  21-DECEMBER-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/117,991
;      FILING DATE:  08-SEPTEMBER-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/931,473
;      FILING DATE:  17-AUGUST-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/964,624
;      FILING DATE:  21-OCTOBER-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/536,428
;      FILING DATE:  11-JUNE-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/714,131
;      FILING DATE:  10-JUNE-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/536,428
;      FILING DATE:  11-JUNE-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Barry J. Swanson
;      REGISTRATION NUMBER:  33,215
;      REFERENCE/DOCKET NUMBER:  NEX17/PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 793-3333
;      TELEFAX:  (303) 793-3433
;  INFORMATION FOR SEQ ID NO:  120:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  76 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
PCT-US95-05600-120
```

```
  Query Match             57.3%;  Score 12.6;  DB 5;  Length 76;
  Best Local Similarity   57.9%;  Pred. No. 6.2e+02;
  Matches   11;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       4 CTGTGAACGTTCGAGATGA 22
           |:  ||| |::||| |:||
Db      44 CUUAGAAAGUUCGACAUGA 62


RESULT  45
US-08-447-169A-21
; Sequence 21, Application US/08447169A
; Patent No. 5811533
;  GENERAL INFORMATION:
;    APPLICANT:  JANJIC, N. and GOLD, L.
;    TITLE OF INVENTION:  HIGH-AFFINITY OLIGONUCLEOTIDE
;    TITLE OF INVENTION:  LIGANDS TO VASCULAR ENDOTHELIAL
;    TITLE OF INVENTION:  GROWTH FACTOR (VEGF)
;    NUMBER OF SEQUENCES:  242
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Swanson & Bratschun, L.L.C.
;      STREET:  8400 E. Prentice Place, Suite 200
;      CITY:  Englewood
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.44 MG storage
;      COMPUTER:  IBM compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/447,169A
;      FILING DATE:  19-MAY-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/233,012
;      FILING DATE:  25-APRIL-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/205,515
;      FILING DATE:  03-MARCH-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/964,624
;      FILING DATE:  21-OCTOBER-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/714,131
;      FILING DATE:  10-JUNE-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/536,428
;      FILING DATE:  11-JUNE-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Barry J. Swanson
;      REGISTRATION NUMBER:  33,215
;      REFERENCE/DOCKET NUMBER:  NEX14
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 793-3333
;      TELEFAX:  (303) 793-3433
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  77 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-447-169A-21


  Query Match             57.3%;  Score 12.6;  DB 1;  Length 77;
  Best Local Similarity   57.9%;  Pred. No. 6.2e+02;
  Matches   11;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       4 CTGTGAACGTTCGAGATGA 22
           | |:|  ||::||| |:||
Db      45 CGGUGCGCGUUCGACAUGA 63
```

Search completed: November 29, 2001, 14:48:19
Job time: 3592 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: November 29, 2001, 14:23:50 ; Search time 1878.42 Seconds
(without alignments)
125.854 Million cell updates/sec

Title: SEQ1
Perfect score: 22
Sequence: 1 TGACTGTGAACGTTCGAGATGA 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 260912

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_htc:*
10: gb_est1:*
11: gb_est2:*
12: gb_htc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		15.6	70.9	63	13	AZ431742	AZ431742 1M0216O18
2		14.8	67.3	99	10	AI313875	AI313875 SWOvAFCAP
c 3		14	63.6	67	10	AA748429	AA748429 ny01b05.s
c 4		14	63.6	88	13	AZ583456	AZ583456 1M0378G03
5		13.6	61.8	58	10	AA840471	AA840471 vw76e10.r
6		13.2	60.0	40	10	AA779179	AA779179 zj43c07.s
c 7		13.2	60.0	50	10	AU106360	AU106360 AU106360
c 8		13.2	60.0	100	13	AZ390824	AZ390824 1M0152P20
9		13	59.1	61	10	AA836207	AA836207 od22h05.s
10		13	59.1	62	13	BH127397	BH127397 G-1c17.r
c 11		13	59.1	68	10	AA104737	AA104737 mo50c09.r
c 12		13	59.1	75	13	AZ775986	AZ775986 2M0009J16
c 13		13	59.1	79	11	BF647619	BF647619 NF012E12E
14		13	59.1	86	11	BF506962	BF506962 11349P-9c
15		12.8	58.2	88	13	AZ783178	AZ783178 2M0024H07
c 16		12.6	57.3	29	13	AZ760190	AZ760190 1M0553P09
17		12.6	57.3	57	13	AZ921432	AZ921432 1006030A0
18		12.6	57.3	83	13	AZ590182	AZ590182 1M0399D09
c 19		12.4	56.4	60	10	AA761865	AA761865 nz64b03.s
c 20		12.4	56.4	63	11	BF633413	BF633413 NF055G12D
21		12.4	56.4	65	13	AZ975641	AZ975641 2M0250I16
c 22		12.4	56.4	79	11	BI175653	BI175653 OSTR051F3
c 23		12.4	56.4	82	10	AA406148	AA406148 zu20c11.s
c 24		12.4	56.4	82	13	AZ767894	AZ767894 1M0567M03
25		12.4	56.4	85	10	AA689791	AA689791 vs07h08.r
c 26		12.4	56.4	88	10	AI953694	AI953694 wq47c06.x
27		12.4	56.4	88	10	AJ283191	AJ283191 4A3A-P7F1
28		12.4	56.4	88	13	AZ586476	AZ586476 1M0392J24
29		12.4	56.4	96	10	AA396017	AA396017 vb42e08.r
c 30		12.2	55.5	26	13	AZ352012	AZ352012 1M0090M13
c 31		12.2	55.5	28	13	AZ776616	AZ776616 2M0010K24
c 32		12.2	55.5	32	13	AZ320254	AZ320254 1M0040P07
33		12.2	55.5	48	11	R59822	R59822 yh11d05.r1
34		12.2	55.5	92	10	AA424991	AA424991 zw03h11.r
35		12.2	55.5	92	10	AA509238	AA509238 MBAFCX8H1
c 36		12.2	55.5	96	13	AZ402172	AZ402172 1M0169B24
c 37		12	54.5	49	10	AI267734	AI267734 ap62g07.x
c 38		12	54.5	50	10	AU107883	AU107883 AU107883
c 39		12	54.5	50	13	TA346C05Q	AL496230 T. brucei
c 40		12	54.5	56	13	AZ492525	AZ492525 1M0326M09
c 41		12	54.5	58	10	AA106075	AA106075 m187e04.r
42		12	54.5	64	11	BG514647	BG514647 dad63b05.
c 43		12	54.5	66	13	TA111H12Q	AL460702 T. brucei
c 44		12	54.5	68	13	CNS03NSG	AL252457 Tetraodon
c 45		12	54.5	71	13	D86873	D86873 Human exon

ALIGNMENTS

RESULT 1
AZ431742
LOCUS AZ431742 63 bp DNA GSS 03-OCT-2000
DEFINITION 1M0216O18R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0216O18 R, DNA sequence.
ACCESSION AZ431742
VERSION AZ431742.1 GI:10555755
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 63)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0216 row: O column: 18
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 63.
FEATURES Location/Qualifiers
source 1. .63

```
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0216O18"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       16 a      8 c     14 g     25 t
ORIGIN


  Query Match             70.9%;  Score 15.6;  DB 13;  Length 63;
  Best Local Similarity   81.8%;  Pred. No. 1.2e+03;
  Matches   18;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATGA 22
           ||| |||||| ||| || ||||
Db      36 TGAATGTGAATGTTTGAAATGA 57


RESULT   2
AI313875
LOCUS       AI313875       99 bp    mRNA            EST       16-DEC-1998
DEFINITION  SWOvAFCAP17G08SK Onchocerca volvulus adult female cDNA
            (SAW98MLW-OvAF) Onchocerca volvulus cDNA clone SWOvAFCAP17G08 5',
            mRNA sequence.
ACCESSION   AI313875
VERSION     AI313875.1  GI:4028863
KEYWORDS    EST.
SOURCE      Onchocerca volvulus.
  ORGANISM  Onchocerca volvulus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
            Onchocercidae; Onchocerca.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Lizotte-Waniewski,M. and Williams,S.A.
  TITLE     Genes expressed in adult female stage of Onchocerca volvulus
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Steven A. Williams
            Molecular Parasitology
            Smith College Department of Biological Sciences
            Department of Biological Sciences, Clark Science Center, Smith
            College, Northampton, MA, 01063, USA
            Tel: 4135853826
            Fax: 4135853786
            Email: genome@smith.edu
            Seq primer: pBluescript SK.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Onchocerca volvulus"
                     /db_xref="taxon:6282"
                     /clone="SWOvAFCAP17G08"
                     /clone_lib="Onchocerca volvulus adult female cDNA
                     (SAW98MLW-OvAF)"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="XL1-Blue MRF'"
                     /note="Vector: Lambda Uni-ZAP XR; Site_1: Eco RI; Site_2:
                     Xho I; Filarial nematode parasite of humans. Two adult
                     female worms of  Onchocerca volvulus were isolated from
                     consenting patients and quick frozen.  Adult female mRNA
                     was converted to double-stranded cDNA using reverse
                     transcriptase and oligo(dT) followed by RNase H and DNA
                     pol I.  The library has 7 x 10E5 independent recombinants
                     and the average    insert size is ~1100bp.The library was
                     constructed by Michelle  Lizotte-Waniewski with worms
                     provided by Dr. Sara Lustigman. The library is  available
                     from Dr.Steven A. Williams, email: genome@smith.edu."
BASE COUNT       46 a     10 c     14 g     25 t      4 others
ORIGIN


  Query Match             67.3%;  Score 14.8;  DB 10;  Length 99;
  Best Local Similarity   88.9%;  Pred. No. 3.2e+03;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 TGTGAACGTTCGAGATGA 22
           |||||||||  || |||||
Db       8 TGTGAACGTACGTGATGA 25


RESULT   3
AA748429/c
LOCUS       AA748429       67 bp    mRNA            EST       18-FEB-1998
DEFINITION  ny01b05.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1270449 3',
            mRNA sequence.
ACCESSION   AA748429
VERSION     AA748429.1  GI:2788387
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 67)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
            Ph.D., Gerald Marti, M.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 863    Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 55.
FEATURES             Location/Qualifiers
     source          1. .67
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1270449"
                     /clone_lib="NCI_CGAP_GCB1"
                     /tissue_type="germinal center B cell"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from human tonsillar cells enriched for
                     germinal center B cells by flow sorting (CD20+, IgD-),
                     provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
                     (NCI) and Dr. Gerald Marti (CBER).  cDNA synthesis was
                     primed with a Not I - oligo(dT) primer
                     [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCCTCATTTTTTTTTTTTTTTTT-3'
```

]. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 15 a 13 c 11 g 28 t
ORIGIN

Query Match 63.6%; Score 14; DB 10; Length 67;
Best Local Similarity 77.3%; Pred. No. 6.9e+03;
Matches 17; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy      1 TGACTGTGAACGTTCGAGATGA 22
          || || |||| ||  |||||||
Db     49 TGCCTTTGAAAGTGGGAGATGA 28
```

RESULT 4
AZ583456/c
LOCUS AZ583456 88 bp DNA GSS 13-DEC-2000
DEFINITION 1M0378G03F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0378G03 F, DNA sequence.
ACCESSION AZ583456
VERSION AZ583456.1 GI:11703357
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 88)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0378 row: G column: 03
Seq primer: CGTTGTAAAACGACGGCCAGT
Class: plasmid ends
High quality sequence stop: 88.
FEATURES Location/Qualifiers
source 1. .88
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0378G03"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT 18 a 22 c 21 g 27 t
ORIGIN

Query Match 63.6%; Score 14; DB 13; Length 88;
Best Local Similarity 77.3%; Pred. No. 7.5e+03;
Matches 17; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy      1 TGACTGTGAACGTTCGAGATGA 22
          ||| ||||| || ||||| ||
Db     54 TGAATGTGAGTGTACGAGACGA 33
```

RESULT 5
AA840471
LOCUS AA840471 58 bp mRNA EST 27-FEB-1998
DEFINITION vw76e10.r1 Stratagene mouse heart (#937316) Mus musculus cDNA clone IMAGE:1260906 5', mRNA sequence.
ACCESSION AA840471
VERSION AA840471.1 GI:2916130
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 58)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:663458
Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 52.
FEATURES Location/Qualifiers
source 1. .58
/organism="Mus musculus"
/strain="NIH/Swiss"
/db_xref="taxon:10090"
/clone="IMAGE:1260906"
/clone_lib="Stratagene mouse heart (#937316)"
/sex="pooled"
/tissue_type="heart"
/dev_stage="13 day embryos"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: heart; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dT. 93 pooled NIH/Swiss 13 day embryo hearts. Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"
BASE COUNT 3 a 17 c 27 g 11 t
ORIGIN

Query Match 61.8%; Score 13.6; DB 10; Length 58;

```
  Best Local Similarity   80.0%;  Pred. No. 1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGAT 20
           |||| |||| |||||| | |
Db      23 TGACCGTGAGCGTTCGTGGT 42


RESULT   6
AA779179
LOCUS       AA779179       40 bp    mRNA            EST       05-FEB-1998
DEFINITION  zj43c07.s1 Soares_fetal_liver_spleen_1NFLS_S1 Homo sapiens cDNA
            clone IMAGE:453036 3' similar to TR:Q13537 Q13537 MER37
            TRANSPOSABLE ELEMENT, COMPLETE CONSENSUS SEQUENCE. ;contains
            MER37.t2 MER37 repetitive element ;, mRNA sequence.
ACCESSION   AA779179
VERSION     AA779179.1  GI:2838510
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 40)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .40
                     /organism="Homo sapiens"
                     /db_xref="GDB:1389392"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:453036"
                     /clone_lib="Soares_fetal_liver_spleen_1NFLS_S1"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     This is a subtracted version of the original Soares fetal
                     liver spleen 1NFLS library.  1st strand cDNA was primed
                     with a Pac I - oligo(dT) primer [5'
                     AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT        9 a     10 c      7 g     14 t
ORIGIN


  Query Match             60.0%;  Score 13.2;  DB 10;  Length 40;
  Best Local Similarity   83.3%;  Pred. No. 1.5e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       4 CTGTGAACGTTCGAGATG 21
           ||||||| || | |||||
Db       7 CTGTGAAAGTCCTAGATG 24


RESULT   7
AU106360/c
LOCUS       AU106360       50 bp    mRNA            EST       05-APR-2001
DEFINITION  AU106360 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HEP16401, mRNA sequence.
ACCESSION   AU106360
VERSION     AU106360.1  GI:13555881
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
            ,K., Suyama,A. and Sugano,S.
  TITLE     Fine Structural analysis of transcription start sites of human
            mRNAs using full-length enriched and 5'-end enriched cDNA libraries
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HEP16401"
                     /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT        4 a     19 c      12 g     15 t
ORIGIN


  Query Match             60.0%;  Score 13.2;  DB 10;  Length 50;
  Best Local Similarity   83.3%;  Pred. No. 1.6e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GACTGTGAACGTTCGAGA 19
           ||| |||||||| ||| | .
Db      37 GACGGTGAACGTACGACA 20


RESULT   8
AZ390824/c
LOCUS       AZ390824       100 bp    DNA             GSS       03-OCT-2000
DEFINITION  1M0152P20F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0152P20 F, DNA sequence.
ACCESSION   AZ390824
VERSION     AZ390824.1  GI:10505867
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 100)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
```

```
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0152  row: P  column: 20
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 100.
FEATURES             Location/Qualifiers
     source          1. .100
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0152P20"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       17 a      17 c      47 g      19 t
ORIGIN


  Query Match             60.0%;  Score 13.2;  DB 13;  Length 100;
  Best Local Similarity   83.3%;  Pred. No. 1.9e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAG 18
          |||||| ||| |||| ||
Db     38 TGACTGGGAAAGTTCCAG 21


RESULT  9
AA836207
LOCUS       AA836207       61 bp    mRNA            EST       25-MAR-1998
DEFINITION  od22h05.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1368729
            similar to TR:Q92931 Q92931 3-HYDROXYISOBUTYRYL-COENZYME A
            HYDROLASE. ;, mRNA sequence.
ACCESSION   AA836207
VERSION     AA836207.1  GI:2910526
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 61)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
            Ph.D., Gerald Marti, M.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Insert Length: 872   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .61
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1368729"
                     /clone_lib="NCI_CGAP_GCB1"
                     /tissue_type="germinal center B cell"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from human tonsillar cells enriched for
                     germinal center B cells by flow sorting (CD20+, IgD-),
                     provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
                     (NCI) and Dr. Gerald Marti (CBER).  cDNA synthesis was
                     primed with a Not I - oligo(dT) primer
                     [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCCTCATTTTTTTTTTTTTTTTT-3'
                     ]. Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT       13 a      7 c      17 g      23 t       1 others
ORIGIN


  Query Match             59.1%;  Score 13;  DB 10;  Length 61;
  Best Local Similarity   76.2%;  Pred. No. 2.1e+04;
  Matches   16;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATG 21
          ||| |||||   || ||||||
Db      1 TGAGTGTGATGTTTAGAGATG 21


RESULT  10
BH127397
LOCUS       BH127397       62 bp    DNA             GSS       23-JUL-2001
DEFINITION  G-1c17.r Maize Random Small-insert Genomic Library Zea mays genomic
            clone G-1c17 both, DNA sequence.
ACCESSION   BH127397
VERSION     BH127397.1  GI:14995229
KEYWORDS    GSS.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 62)
  AUTHORS   Meyers,B.C., Tingey,S.V. and Morgante,M.
  TITLE     Abundance, distribution and transcriptional activity of repetitive
            elements in the maize genome
  JOURNAL   Genome Res. (2001) In press
COMMENT     Contact: Morgante M
            Suite 200
            Dupont Genomics
            PO Box 6104, Newark, DE 19714-6104, USA
            Tel: 302 631 2638
            Fax: 302 631 2607
            Email: Michele.morgante@usa.dupont.com
            Sequences were trimmed to include only high quality bases; forward
            and reverse reads were assembled when significant overlaps were
            detected.
            Seq primer: M13reverse
```

```
          Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .62
                     /organism="Zea mays"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="G-1c17"
                     /clone_lib="Maize Random Small-insert Genomic Library"
                     /sex="hermaphrodite"
                     /tissue_type="leaf"
                     /cell_type="Young leaf"
                     /dev_stage="seedling"
                     /note="Vector: pCR-Script; Total genomic DNA was nebulized
                     ; ends were polished with Pfu polymerase and the fragments
                     cloned into pCR-Script."
BASE COUNT       12 a     14 c     19 g     17 t
ORIGIN


  Query Match             59.1%;  Score 13;  DB 13;  Length 62;
  Best Local Similarity   76.2%;  Pred. No. 2.1e+04;
  Matches   16;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       2 GACTGTGAACGTTCGAGATGA 22
           || || || || | |||||||
Db      26 GAATGAGACCGATGGAGATGA 46


RESULT  11
AA104737/c
LOCUS       AA104737       68 bp    mRNA            EST       29-OCT-1996
DEFINITION  mo50c09.r1 Life Tech mouse embryo 10 5dpc 10665016 Mus musculus
            cDNA clone IMAGE:557008 5', mRNA sequence.
ACCESSION   AA104737
VERSION     AA104737.1  GI:1650951
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 68)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:337800
            Seq primer: -28M13 rev1 from Amersham
            High quality sequence stop: 60.
FEATURES             Location/Qualifiers
     source          1. .68
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:557008"
                     /clone_lib="Life Tech mouse embryo 10 5dpc 10665016"
                     /tissue_type="embryo"
                     /dev_stage="10.5dpc embryos"
                     /lab_host="DH10B"
                     /note="Organ: whole embryo; Vector: pCMV-SPORT2; Site_1:
                     SalI; Site_2: NotI; Cloned unidirectionally.  Primer:
                     Oligo dT.  10.5dpc embryos. pCMV-SPORT2 vector."
BASE COUNT       23 a     22 c     11 g     12 t
ORIGIN


  Query Match             59.1%;  Score 13;  DB 10;  Length 68;
  Best Local Similarity   76.2%;  Pred. No. 2.1e+04;
  Matches   16;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGATG 21
           |||||  ||||||||  || |
Db      68 TGACTTGGAACGTTCCCGAGG 48


RESULT  12
AZ775986/c
LOCUS       AZ775986       75 bp    DNA             GSS       16-FEB-2001
DEFINITION  2M0009J16F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0009J16 F, DNA sequence.
ACCESSION   AZ775986
VERSION     AZ775986.1  GI:12903094
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 75)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0009  row: J  column: 16
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 75.
FEATURES             Location/Qualifiers
     source          1. .75
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0009J16"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
```

```
                and selected for ampicillin resistance."
BASE COUNT       21 a     16 c     18 g     20 t
ORIGIN


  Query Match             59.1%;  Score 13;  DB 13;  Length 75;
  Best Local Similarity   76.2%;  Pred. No. 2.2e+04;
  Matches   16;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATG 21
          ||| ||||||  || |||| |
Db     55 TGAATGTGAATTTTGGAGAAG 35


RESULT  13
BF647619/c
LOCUS       BF647619       79 bp    mRNA            EST       20-DEC-2000
DEFINITION  NF012E12EC1F1097 Elicited cell culture Medicago truncatula cDNA
            clone NF012E12EC 5', mRNA sequence.
ACCESSION   BF647619
VERSION     BF647619.1  GI:11912761
KEYWORDS    EST.
SOURCE      barrel medic.
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 79)
  AUTHORS   Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
            Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.
  TITLE     Expressed Sequence Tags from the Samuel Roberts Noble Foundation -
            Center for Medicago Genomics Research
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Dixon RA
            Plant Biology Division
            The Samuel Roberts Noble Foundation
            2510 Sam Noble Parkway, Ardmore, OK 73402, USA
            Tel: 580 221 7302
            Fax: 580 221 7380
            Email: radixon@noble.org
            Insert Length: 79   Std Error: 0.00
            Plate: 012  row: E  column: 12
            Seq primer: TCACACAGGAAACAGCTATGAC.
FEATURES             Location/Qualifiers
     source          1. .79
                     /organism="Medicago truncatula"
                     /db_xref="taxon:3880"
                     /clone="NF012E12EC"
                     /clone_lib="Elicited cell culture"
                     /tissue_type="Cell cultures derived from root tissues"
                     /dev_stage="Cell suspensions were subcultured every 14
                     days. Cells were induced six days after subculture"
                     /note="Vector: Lambda Zap; Cells were induced with yeast
                     cell wall extracts equivalent to  50ug/ml glucose in the
                     final concentration. Samples were taken at  0.5, 1, 12 and
                     24 hours after induction. Equal amounts of RNA from  each
                     time point were pooled and used for mRNA isolation."
BASE COUNT       32 a     20 c     10 g     14 t      3 others
ORIGIN


  Query Match             59.1%;  Score 13;  DB 11;  Length 79;
  Best Local Similarity   76.2%;  Pred. No. 2.2e+04;
  Matches   16;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      2 GACTGTGAACGTTCGAGATGA 22
          || | |||| |||| |||| |
Db     62 GAGTTTGAAGGTTCTAGATTA 42


RESULT  14
BF506962
LOCUS       BF506962       86 bp    mRNA            EST       07-DEC-2000
DEFINITION  11349P-9c Pooled green leaf and root tissue Sorghum bicolor cDNA
            clone 11349P-9, mRNA sequence.
ACCESSION   BF506962
VERSION     BF506962.1  GI:11590260
KEYWORDS    EST.
SOURCE      sorghum.
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 86)
  AUTHORS   Childs,K.L., Klein,R.R., Klein,P.E., Morishige,D.T. and Mullet,J.E.
  TITLE     Mapping Genes on an Integrated Sorghum Genetic and Physical Map
            Using cDNA Selection Technology
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Kevin Childs
            Department of Biochemistry and Biophysics
            Texas A&M University
            College Station, TX 77843, USA
            Tel: 979 845 0832
            Fax: 979 862 4718
            Email: kchilds@unix.tamu.edu.
FEATURES             Location/Qualifiers
     source          1. .86
                     /organism="Sorghum bicolor"
                     /cultivar="BTx623"
                     /db_xref="taxon:4558"
                     /clone="11349P-9"
                     /clone_lib="Pooled green leaf and root tissue"
                     /tissue_type="green leaf and root tissue"
                     /note="Vector: pBluescript II (SK); Site_1: EcoRI; Site_2:
                     EcoRI"
BASE COUNT       18 a     15 c     23 g     30 t
ORIGIN


  Query Match             59.1%;  Score 13;  DB 11;  Length 86;
  Best Local Similarity   76.2%;  Pred. No. 2.2e+04;
  Matches   16;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATG 21
          ||||  |||  ||||||| ||
Db     25 TGACGTTGATTGTTCGAGTTG 45


RESULT  15
AZ783178
LOCUS       AZ783178       88 bp    DNA             GSS       16-FEB-2001
DEFINITION  2M0024H07R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0024H07 R, DNA sequence.
ACCESSION   AZ783178
VERSION     AZ783178.1  GI:12917646
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 88)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
```

```
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0024  row: H  column: 07
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 88.
FEATURES              Location/Qualifiers
     source           1. .88
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="UUGC2M0024H07"
                      /clone_lib="Mouse 10kb plasmid UUGC1M library"
                      /sex="Male"
                      /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                      /note="Vector: PWD42nv; Purified genomic DNA from M.
                      musculus C57BL/6J (male) was obtained  from the Jackson
                      Laboratory Mouse DNA Resource
                      (http://www.jax.org/resources/documents/dnares/). The DNA
                      was hydrodynamically sheared by repeated passage through a
                      0.005 inch orifice at constant velocity. The sheared DNA
                      was blunt end-repaired with T4 DNA polymerase and T4
                      polynucleotide kinase. Adaptor oligonucleotides were
                      ligated to the blunt ends in high  molar excess. The
                      adaptored DNA was purified and size-selected for a 9.5 to
                      10.5 kb range using preparative agarose gel
                      electrophoresis. Vector DNA was prepared from a derivative
                      of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                      inducible derivative of plasmid R1. The vector was ligated
                      with adaptors complementary to the insert adaptors and
                      purified. The sheared, adaptored mouse DNA was annealed to
                      adaptored vector DNA, and transformed into
                      chemically-competent E. coli XL10-Gold (Stratagene) cells
                      and selected for ampicillin resistance."
BASE COUNT       16 a     24 c     18 g     30 t
ORIGIN


  Query Match             58.2%;  Score 12.8;  DB 13;  Length 88;
  Best Local Similarity   87.5%;  Pred. No. 2.8e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCG 16
          |||||||||||  || |
Db     66 TGACTGTGAACATTAG 81


RESULT  16
AZ760190/c
LOCUS       AZ760190       29 bp    DNA             GSS       16-FEB-2001
DEFINITION  1M0553P09R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0553P09 R, DNA sequence.
ACCESSION   AZ760190
VERSION     AZ760190.1  GI:12867754
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 29)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0553  row: P  column: 09
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 29.
FEATURES              Location/Qualifiers
     source           1. .29
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="UUGC1M0553P09"
                      /clone_lib="Mouse 10kb plasmid UUGC1M library"
                      /sex="Male"
                      /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                      /note="Vector: PWD42nv; Purified genomic DNA from M.
                      musculus C57BL/6J (male) was obtained  from the Jackson
                      Laboratory Mouse DNA Resource
                      (http://www.jax.org/resources/documents/dnares/). The DNA
                      was hydrodynamically sheared by repeated passage through a
                      0.005 inch orifice at constant velocity. The sheared DNA
                      was blunt end-repaired with T4 DNA polymerase and T4
                      polynucleotide kinase. Adaptor oligonucleotides were
                      ligated to the blunt ends in high  molar excess. The
                      adaptored DNA was purified and size-selected for a 9.5 to
                      10.5 kb range using preparative agarose gel
                      electrophoresis. Vector DNA was prepared from a derivative
                      of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                      inducible derivative of plasmid R1. The vector was ligated
                      with adaptors complementary to the insert adaptors and
                      purified. The sheared, adaptored mouse DNA was annealed to
                      adaptored vector DNA, and transformed into
                      chemically-competent E. coli XL10-Gold (Stratagene) cells
                      and selected for ampicillin resistance."
BASE COUNT        5 a     11 c      8 g      5 t
ORIGIN


  Query Match             57.3%;  Score 12.6;  DB 13;  Length 29;
  Best Local Similarity   78.9%;  Pred. No. 2.6e+04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGA 19
          |||||||||  ||  || ||
Db     24 TGACTGTGACTGTGCGGGA 6


RESULT  17
AZ921432
LOCUS       AZ921432       57 bp    DNA             GSS       20-MAR-2001
DEFINITION  1006030A03.2EL_x2 1006 - RescueMu Grid G Zea mays genomic, DNA
            sequence.
ACCESSION   AZ921432
VERSION     AZ921432.1  GI:13393068
KEYWORDS    GSS.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 57)
  AUTHORS   Walbot,V.
  TITLE     Maize genomic sequences found using engineered RescueMu transposon
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Possible ligation site of ends cut by 2 different endonucleases.
```

```
          Reverse complemented post-ligation sequence from source sequence.
          Plate: 1006030  row: A  column: 03
          Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .57
                     /organism="Zea mays"
                     /cultivar="mixed background W23/A188/B73"
                     /db_xref="taxon:4577"
                     /clone_lib="1006 - RescueMu Grid G"
                     /tissue_type="leaf"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: leaf; Vector: RescueMu (engineered from
                     pBlueScript backbone); Site_1: BamHI; Site_2: BglII;
                     RescueMu is a 4.9 kb, modified maize Mu transposon
                     designed to allow plasmid rescue from total genomic DNA.
                     Mu elements insert preferentially into transcription
                     units. For more information on RescueMu, go to the web
                     site 'www.zmdb.iastate.edu' and follow the links for
                     'RescueMu.' Grid G was grown at Stanford in 2000. DNA was
                     extracted from leaf punches, double digested using BamHI
                     and BglII, and ligated to form circular plasmids. DH10B
                     cells were transformed and then screened on LB plates with
                     ampicillin."
BASE COUNT       20 a      7 c     14 g     16 t
ORIGIN


  Query Match             57.3%;  Score 12.6;  DB 13;  Length 57;
  Best Local Similarity   78.9%;  Pred. No. 3.1e+04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      4 CTGTGAACGTTCGAGATGA 22
          |||| |  |||||||| ||
Db     27 CTGTTATGGTTCGAGAAGA 45


RESULT  18
AZ590182
LOCUS       AZ590182       83 bp    DNA             GSS       13-DEC-2000
DEFINITION  1M0399D09R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0399D09 R, DNA sequence.
ACCESSION   AZ590182
VERSION     AZ590182.1  GI:11712372
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 83)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0399  row: D  column: 09
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 83.
FEATURES             Location/Qualifiers
     source          1. .83
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0399D09"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       22 a      23 c     13 g     25 t
ORIGIN


  Query Match             57.3%;  Score 12.6;  DB 13;  Length 83;
  Best Local Similarity   78.9%;  Pred. No. 3.5e+04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 GACTGTGAACGTTCGAGAT 20
          ||||||||||  |||  ||
Db     25 GACTGTGAACTGTCGGCAT 43


RESULT  19
AA761865/c
LOCUS       AA761865       60 bp    mRNA            EST       07-FEB-1998
DEFINITION  nz64b03.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1300205 3',
            mRNA sequence.
ACCESSION   AA761865
VERSION     AA761865.1  GI:2810795
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
            Ph.D., Gerald Marti, M.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 726   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 52.
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
```

```
                     /clone="IMAGE:1300205"
                     /clone_lib="NCI_CGAP_GCB1"
                     /tissue_type="germinal center B cell"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from human tonsillar cells enriched for
                     germinal center B cells by flow sorting (CD20+, IgD-),
                     provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
                     (NCI) and Dr. Gerald Marti (CBER).  cDNA synthesis was
                     primed with a Not I - oligo(dT) primer
                     [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCCTCATTTTTTTTTTTTTTTTTT-3'
                     ]. Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT       13 a      7 c      6 g     34 t
ORIGIN


  Query Match             56.4%;  Score 12.4;  DB 10;  Length 60;
  Best Local Similarity   72.7%;  Pred. No. 4e+04;
  Matches   16;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          ||||  | |||  | |||||||
Db     41 TGACAATAAACCATAGAGATGA 20


RESULT  20
BF633413/c
LOCUS       BF633413       63 bp    mRNA            EST       19-DEC-2000
DEFINITION  NF055G12DT1F1100 Drought Medicago truncatula cDNA clone NF055G12DT
            5', mRNA sequence.
ACCESSION   BF633413
VERSION     BF633413.1  GI:11897571
KEYWORDS    EST.
SOURCE      barrel medic.
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 63)
  AUTHORS   Torrez-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
            Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.
  TITLE     Expressed Sequence Tags from the Samuel Roberts Noble Foundation
            Medicago truncatula drought library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: May GD
            Plant Biology Division
            The Samuel Roberts Noble Foundation
            2510 Sam Noble Parkway, Ardmore, OK 73402, USA
            Tel: 580 221 7391
            Fax: 580 221 7380
            Email: gdmay@noble.org
            Insert Length: 63   Std Error: 0.00
            Plate: 055  row: G  column: 12
            Seq primer: TCACACAGGAAACAGCTATGAC.
FEATURES             Location/Qualifiers
     source          1. .63
                     /organism="Medicago truncatula"
                     /db_xref="taxon:3880"
                     /clone="NF055G12DT"
                     /clone_lib="Drought"
                     /tissue_type="Plantlets"
                     /dev_stage="Pooled timepoints"
                     /note="Vector: Lambda Zap; Contains a mixture of entire
                     plantlets harvested in a series of days-post-watering
                     timepoints. "
BASE COUNT       21 a     22 c      3 g     17 t
ORIGIN


  Query Match             56.4%;  Score 12.4;  DB 11;  Length 63;
  Best Local Similarity   72.7%;  Pred. No. 4e+04;
  Matches   16;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          ||| | ||||  |  |||||||
Db     56 TGATTATGAAATTGAGAGATGA 35


RESULT  21
AZ975641
LOCUS       AZ975641       65 bp    DNA             GSS       27-APR-2001
DEFINITION  2M0250I16R Mouse 10kb plasmid UUGC2M library Mus musculus genomic
            clone UUGC2M0250I16 R, DNA sequence.
ACCESSION   AZ975641
VERSION     AZ975641.1  GI:13846868
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 65)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0250  row: I  column: 16
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 65.
FEATURES             Location/Qualifiers
     source          1. .65
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0250I16"
                     /clone_lib="Mouse 10kb plasmid UUGC2M library"
                     /sex="Female"
                     /lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (female) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       7 a      1 c     31 g     26 t
```

```
ORIGIN


  Query Match             56.4%;  Score 12.4;  DB 13;  Length 65;
  Best Local Similarity   72.7%;  Pred. No. 4.1e+04;
  Matches   16;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          ||  |||||| ||  | |||||
Db     37 TGTGTGTGAAGGTATGGGATGA 58


RESULT  22
BI175653/c
LOCUS       BI175653       79 bp    mRNA            EST       09-JUL-2001
DEFINITION  OSTR051F3_1 AD-wrmcDNA Caenorhabditis elegans cDNA similar to
            C14B1.2, mRNA sequence.
ACCESSION   BI175653
VERSION     BI175653.1  GI:14641456
KEYWORDS    EST.
SOURCE      Caenorhabditis elegans.
  ORGANISM  Caenorhabditis elegans
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
            ; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE   1  (bases 1 to 79)
  AUTHORS   Reboul,J., Vaglio,P., Tzellas,N., Thierry-Mieg,N., Moore,T.,
            Jackson,C., Shin-i,T., Kohara,Y., Thierry-Mieg,D., Thierry-Mieg,J.,
            Lee,H., Hitti,J., Doucette-Stamm,L., Hartley,J.L., Temple,G.F.,
            Brasch,M.A., Vandenhaute,J., Lamesch,P.E., Hill,D.E. and Vidal,M.
  TITLE     Open-reading-frame sequence tags (OSTs) support the existence of at
            least 17,300 genes in C. elegans
  JOURNAL   Nat. Genet. 27 (3), 332-336 (2001)
  MEDLINE   21135099
COMMENT     Contact: Reboul J, Vaglio P
            Marc Vidal Laboratory
            Dana Farber Cancer Institute
            44 Binney Street, Boston, MA 02115, USA
            Tel: 617 632 5180
            Fax: 617 632 2425
            Email: Jerome_Reboul@dfci.harvard.edu
            Sequence tag of Gateway entry clones. The primers used were
            designed on the predicted protein encoding ORF. C. elegans ORFeome
            cloning project : Contact jerome_reboul@dfci.harvard.edu or
            philippe_vaglio@dfci.harvard.edu
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .79
                     /organism="Caenorhabditis elegans"
                     /strain="N2"
                     /db_xref="taxon:6239"
                     /clone_lib="AD-wrmcDNA"
                     /sex="Hermaphrodite and male"
                     /tissue_type="whole animal"
                     /dev_stage="mixed stage"
                     /note="The AD-wrmcDNA library was generated with poly(A)+
                     RNA isolated from both hermaphrodite and male N2 worms of
                     all larval stages, embryos, adults and dauers and the
                     subsequent generation of cDNAs by poly(A) priming. The
                     cDNAs were cloned into pPC86"
BASE COUNT       19 a     22 c     17 g     21 t
ORIGIN


  Query Match             56.4%;  Score 12.4;  DB 11;  Length 79;
  Best Local Similarity   72.7%;  Pred. No. 4.3e+04;
  Matches   16;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          || || || | || ||||| ||
Db     63 TGGCTTTGGAAGTCCGAGAGGA 42


RESULT  23
AA406148/c
LOCUS       AA406148       82 bp    mRNA            EST       17-MAY-1997
DEFINITION  zu20c11.s1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:738548 3'
            similar to TR:E211619 E211619 STOP PROTEIN. ;, mRNA sequence.
ACCESSION   AA406148
VERSION     AA406148.1  GI:2064129
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 82)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.
            , Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie
            ,T., Waterston,R. and Wilson,R.
  TITLE     WashU-Merck EST Project 1997
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -41m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .82
                     /organism="Homo sapiens"
                     /db_xref="GDB:5946214"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:738548"
                     /clone_lib="Soares_NhHMPu_S1"
                     /tissue_type="Pooled human melanocyte, fetal heart, and
                     pregnant uterus"
                     /lab_host="DH10B"
                     /note="Organ: mixed (see below); Vector: pT7T3D-Pac
                     (Pharmacia) with a modified polylinker; Site_1: Not I;
                     Site_2: Eco RI; Equal amounts of plasmid DNA from three
                     normalized libraries (melanocyte 2NbHM, pregnant uterus
                     NbHPU, and fetal heart NbHH19W) were mixed, and ss circles
                     were made in vitro. Following HAP purification, this DNA
                     was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from pools of
                     5,000 clones made from the same 3 libraries. The pools
                     consisted of I.M.A.G.E. clones 260232-265223,
                     340488-345479, and 484488-489479."
BASE COUNT       20 a     20 c     22 g     20 t
ORIGIN


  Query Match             56.4%;  Score 12.4;  DB 10;  Length 82;
  Best Local Similarity   72.7%;  Pred. No. 4.3e+04;
  Matches   16;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          || || |||  | || ||||||
Db     39 TGCCTCTGATGGATCAAGATGA 18


RESULT  24
AZ767894/c
LOCUS       AZ767894       82 bp    DNA             GSS       16-FEB-2001
DEFINITION  1M0567M03R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0567M03 R, DNA sequence.
ACCESSION   AZ767894
VERSION     AZ767894.1  GI:12886458
KEYWORDS    GSS.
SOURCE      house mouse.
```

```
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 82)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0567  row: M  column: 03
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 82.
FEATURES             Location/Qualifiers
     source          1. .82
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0567M03"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       18 a     16 c     26 g     22 t
ORIGIN


  Query Match             56.4%;  Score 12.4;  DB 13;  Length 82;
  Best Local Similarity   72.7%;  Pred. No. 4.3e+04;
  Matches   16;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          || ||||| | | || || |||
Db     56 TGGCTGTGCAGGCTCCAGCTGA 35


RESULT  25
AA689791
LOCUS       AA689791       85 bp    mRNA            EST       16-DEC-1997
DEFINITION  vs07h08.r1 Barstead mouse irradiated colon MPLRB7 Mus musculus cDNA
            clone IMAGE:1137567 5' similar to WP:B0361.8 CE00837 ;, mRNA
            sequence.
ACCESSION   AA689791
VERSION     AA689791.1  GI:2690727
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 85)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:618839
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .85
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1137567"
                     /clone_lib="Barstead mouse irradiated colon MPLRB7"
                     /dev_stage="8 weeks"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: EcoRI; Site_2: NotI; Tissue obtained
                     from 8 week old mouse.  Colon was harvested 72 hours after
                     irradiation with 1400 Gys.  1st strand cDNA was primed
                     with  a Not I - oligo(dT) primer
                     [5'TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI
                     adaptors [AATTCGGATCCTTG], digested with Not I and cloned
                     into the Not I and Eco RI sites of  the modified pT7T3
                     vector. Library constructed by Bob Barstead.  "
BASE COUNT       24 a     21 c     22 g     18 t
ORIGIN


  Query Match             56.4%;  Score 12.4;  DB 10;  Length 85;
  Best Local Similarity   72.7%;  Pred. No. 4.4e+04;
  Matches   16;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          || | | ||| || | ||||||
Db     30 TGCCGGAGAAGGTACTAGATGA 51


RESULT  26
AI953694/c
LOCUS       AI953694       88 bp    mRNA            EST       08-MAR-2000
DEFINITION  wq47c06.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:2474410 3'
            similar to TR:P78477 P78477 TPRDIII. ;, mRNA sequence.
ACCESSION   AI953694
VERSION     AI953694.1  GI:5746004
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 88)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
```

```
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
            R. Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Insert Length: 1614   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .88
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:2474410"
                      /clone_lib="NCI_CGAP_GC6"
                      /tissue_type="pooled germ cell tumors"
                      /lab_host="DH10B"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA
                      from the normalized library NCI_CGAP_GC4 was prepared, and
                      ss circles were made in vitro. Following HAP purification,
                      this DNA was used as tracer in a subtractive hybridization
                      reaction. The driver was PCR-amplified cDNAs from a pool
                      of 5,000 clones made from the same library (cloneIDs
                      1257096-1258631, 1469064-1470983, and 1475592-1476743).
                      Subtraction by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT       22 a     24 c     15 g     27 t
ORIGIN


  Query Match             56.4%;  Score 12.4;  DB 10;  Length 88;
  Best Local Similarity   72.7%;  Pred. No. 4.4e+04;
  Matches   16;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          |||| ||   ||||  ||||||
Db     27 TGACAGTACGCGTTTTAGATGA 6


RESULT  27
AJ283191
LOCUS       AJ283191       88 bp    mRNA            EST       30-JUN-2000
DEFINITION  4A3A-P7F11-R Anopheles gambiae immune competent 4A3A Anopheles
            gambiae cDNA clone 4A3A-P7F11, mRNA sequence.
ACCESSION   AJ283191
VERSION     AJ283191.1  GI:6931070
KEYWORDS    EST.
SOURCE      African malaria mosquito.
  ORGANISM  Anopheles gambiae
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea
            ; Anopheles.
REFERENCE   1  (bases 1 to 88)
  AUTHORS   Dimopoulos,G., Casavant,T.L., Chang,S., Scheetz,T., Roberts,C.,
            Donohue,M., Schultz,J., Benes,V., Bork,P., Ansorge,W., Soares,M.B.
            and Kafatos,F.C.
  TITLE     Anopheles gambiae pilot gene discovery project: identification of
            mosquito innate immunity genes from expressed sequence tags
            generated from immune-competent cell lines
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
  MEDLINE   20300950
COMMENT     Contact: Dimopoulos G
            Fotis C. Kafatos laboratory
            European Molecular Biology Laboratory
            Meyerhofstrasse 1, 69117 Heidelberg, Germany.
FEATURES              Location/Qualifiers
     source           1. .88
                      /organism="Anopheles gambiae"
                      /strain="4A r/r"
                      /db_xref="taxon:7165"
                      /clone="4A3A-P7F11"
                      /clone_lib="Anopheles gambiae immune competent 4A3A"
                      /cell_line="immune competent 4A3A"
                      /lab_host="E. coli DH10B"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
                      forward priming site which reads from the 3' end of the
                      cDNA. The 4A3A is a directionally cloned and normalized
                      cDNA library that was constructed from the 4A3A cell line
                      oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
                      (1996) : Normalization and Subtraction: Two approaches To
                      Facilitate Gene Discovery, Genome Research 6, 791-806."
BASE COUNT       19 a     28 c     17 g     24 t
ORIGIN


  Query Match             56.4%;  Score 12.4;  DB 10;  Length 88;
  Best Local Similarity   72.7%;  Pred. No. 4.4e+04;
  Matches   16;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGATGA 22
          | ||  ||  ||||||||| |||
Db     33 TTACATTGCTCGTTCGAGCTGA 54


RESULT  28
AZ586476
LOCUS       AZ586476       88 bp    DNA             GSS       13-DEC-2000
DEFINITION  1M0392J24F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0392J24 F, DNA sequence.
ACCESSION   AZ586476
VERSION     AZ586476.1  GI:11708666
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 88)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0392  row: J  column: 24
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 88.
FEATURES              Location/Qualifiers
     source           1. .88
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="UUGC1M0392J24"
                      /clone_lib="Mouse 10kb plasmid UUGC1M library"
                      /sex="Male"
```

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT 22 a 21 c 20 g 25 t
ORIGIN

Query Match 56.4%; Score 12.4; DB 13; Length 88;
Best Local Similarity 72.7%; Pred. No. 4.4e+04;
Matches 16; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy      1 TGACTGTGAACGTTCGAGATGA 22
          |||||||| | ||  |||  ||
Db     32 TGACTGTGCATGTGAGAGGGGA 53
```

RESULT 29
AA396017
LOCUS AA396017 96 bp mRNA EST 25-APR-1997
DEFINITION vb42e08.r1 Soares mouse lymph node NbMLN Mus musculus cDNA clone IMAGE:751622 5', mRNA sequence.
ACCESSION AA396017
VERSION AA396017.1 GI:2049086
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 96)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:460606
Seq primer: -28m13 rev2 ET from Amersham.
FEATURES Location/Qualifiers
source 1. .96
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:751622"
/clone_lib="Soares mouse lymph node NbMLN"
/sex="male"
/tissue_type="lymph node"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="Organ: lymph node; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCGATACTTTTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. RNA provided by Dr. Bertrand Jordan. Library constructed and normalized by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 31 a 18 c 23 g 23 t 1 others
ORIGIN

Query Match 56.4%; Score 12.4; DB 10; Length 96;
Best Local Similarity 72.7%; Pred. No. 4.5e+04;
Matches 16; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy      1 TGACTGTGAACGTTCGAGATGA 22
          |||||| |||  | | |||| |
Db     65 TGACTGAGAAGATACTAGATAA 86
```

RESULT 30
AZ352012/c
LOCUS AZ352012 26 bp DNA GSS 29-SEP-2000
DEFINITION 1M0090M13F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0090M13 F, DNA sequence.
ACCESSION AZ352012
VERSION AZ352012.1 GI:10431249
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 26)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0090 row: M column: 13
Seq primer: CGTTGTAAAACGACGGCCAGT
Class: plasmid ends
High quality sequence stop: 26.
FEATURES Location/Qualifiers
source 1. .26
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0090M13"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

```
BASE COUNT        8 a      10 c       2 g       6 t
ORIGIN
```

```
  Query Match             55.5%;  Score 12.2;  DB 13;  Length 26;
  Best Local Similarity   82.4%;  Pred. No. 4e+04;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       5 TGTGAACGTTCGAGATG 21
           |||||  ||| ||||||
Db      25 TGTGAGTGTTGGAGATG 9
```

```
RESULT  31
AZ776616/c
LOCUS       AZ776616       28 bp    DNA             GSS       16-FEB-2001
DEFINITION  2M0010K24F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0010K24 F, DNA sequence.
ACCESSION   AZ776616
VERSION     AZ776616.1  GI:12904394
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 28)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0010  row: K  column: 24
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 28.
FEATURES             Location/Qualifiers
     source          1. .28
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0010K24"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
```

0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

```
BASE COUNT        7 a      12 c       3 g       6 t
ORIGIN
```

```
  Query Match             55.5%;  Score 12.2;  DB 13;  Length 28;
  Best Local Similarity   82.4%;  Pred. No. 4e+04;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       6 GTGAACGTTCGAGATGA 22
             |||  | ||||||||||
Db      26 GTGGGCTTTCGAGATGA 10
```

```
RESULT  32
AZ320254/c
LOCUS       AZ320254       32 bp    DNA             GSS       29-SEP-2000
DEFINITION  1M0040P07F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0040P07 F, DNA sequence.
ACCESSION   AZ320254
VERSION     AZ320254.1  GI:10371848
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 32)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0040  row: P  column: 07
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 32.
FEATURES             Location/Qualifiers
     source          1. .32
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0040P07"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
```

was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT 9 a 11 c 3 g 9 t
ORIGIN

Query Match 55.5%; Score 12.2; DB 13; Length 32;
Best Local Similarity 82.4%; Pred. No. 4.2e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 CTGTGAACGTTCGAGAT 20
||||||| ||| ||| |
Db 29 CTGTGAATGTTGGAGCT 13

RESULT 33
R59822
LOCUS R59822 48 bp mRNA EST 24-MAY-1995
DEFINITION yh11d05.r1 Soares infant brain 1NIB Homo sapiens cDNA clone IMAGE:43095 5' similar to gb|K01562|HUMCRHY1 Human Ro RNA (rRNA);, mRNA sequence.
ACCESSION R59822
VERSION R59822.1 GI:830517
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 48)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence starts: 1
High quality sequence stops: 1
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Seq primer: M13RP1
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .48
/organism="Homo sapiens"
/db_xref="GDB:415636"
/db_xref="taxon:9606"
/clone="IMAGE:43095"
/clone_lib="Soares infant brain 1NIB"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: whole brain; Vector: Lafmid BA; Site_1: Not I; Site_2: Hind III; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Hind III adaptors (Pharmacia), digested with Not I and directionally cloned into the Not I and Hind III sites of the Lafmid BA vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 15 a 13 c 7 g 12 t 1 others
ORIGIN

Query Match 55.5%; Score 12.2; DB 11; Length 48;
Best Local Similarity 82.4%; Pred. No. 4.7e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TGACTGTGAACGTTCGA 17
||||||||||| || |
Db 13 TGACTGTGAACAATCAA 29

RESULT 34
AA424991
LOCUS AA424991 92 bp mRNA EST 16-OCT-1997
DEFINITION zw03h11.r1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:768261 5' similar to gb:X51760 ZINC FINGER PROTEIN ZFP-36 (HUMAN);, mRNA sequence.
ACCESSION AA424991
VERSION AA424991.1 GI:2107060
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 92)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE WashU-Merck EST Project 1997
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .92
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:768261"
/clone_lib="Soares_NhHMPu_S1"
/tissue_type="Pooled human melanocyte, fetal heart, and pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below); Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NbHM, pregnant uterus NbHPU, and fetal heart NbHH19W) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."
BASE COUNT 33 a 17 c 24 g 18 t

```
ORIGIN


  Query Match             55.5%;  Score 12.2;  DB 10;  Length 92;
  Best Local Similarity   82.4%;  Pred. No. 5.6e+04;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGA 17
          | |||||||| || |||
Db     74 TAACTGTGAAAGTCCGA 90


RESULT  35
AA509238
LOCUS       AA509238       92 bp    mRNA            EST       08-JUL-1997
DEFINITION  MBAFCX8H10T3 Brugia malayi adult female cDNA (SAW96MLW-BmAF) Brugia
            malayi cDNA clone AFCX8H10 5', mRNA sequence.
ACCESSION   AA509238
VERSION     AA509238.1  GI:2247115
KEYWORDS    EST.
SOURCE      Brugia malayi.
  ORGANISM  Brugia malayi
            Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
            Onchocercidae; Brugia.
REFERENCE   1  (bases 1 to 92)
  AUTHORS   Blaxter,M.L., Waterfall,M., Daub,J., Lizotte,M., Baron,L. and Jones
            ,S.J.
  TITLE     Genes expressed in adult female Brugia malayi
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Blaxter ML
            Institute of Cell, Animal and Population Biology
            University of Edinburgh
            Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
            3JT, UK.
            Tel: +44 131 650 6760
            Fax: +44 131 670 5450
            Email: mark.blaxter@ed.ac.uk
            The ABI trace of this sequence can be viewed at
            http://www.sanger.ac.uk/brugia/AFC/MBAFCX8H10T3.html This is the
            full sequence of the cDNA clone. The polyA tail has been clipped
            and is excluded from this sequence
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1. .92
                     /organism="Brugia malayi"
                     /db_xref="taxon:6279"
                     /clone="AFCX8H10"
                     /clone_lib="Brugia malayi adult female cDNA (SAW96MLW-BmAF
                     )"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="XL1-Blue MRF'"
                     /note="Vector: Lambda Uni-ZAP XR; Site_1: Eco RI; Site_2:
                     Xho I; Lymphatic filarial nematode parasite of humans.
                     mRNA was prepared from approximately 50 adult females
                     isolated from the  peritoneal cavity of jirds and
                     converted to double-stranded cDNA        using reverse
                     transcriptase and oligo(dT) followed by RNase H and DNA
                     pol I. The library has 5 x 10E6 independent recombinants
                     and the average insert size is ~900bp.The library was
                     constructed by          Michelle Lizotte-Waniewski. The
                     library is available from      Dr.S.A.Williams, email:
                     genome@smith.edu."
BASE COUNT       30 a      13 c      24 g      25 t
ORIGIN


  Query Match             55.5%;  Score 12.2;  DB 10;  Length 92;
  Best Local Similarity   82.4%;  Pred. No. 5.6e+04;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      3 ACTGTGAACGTTCGAGA 19
          ||||| ||| ||| |||


Db     42 ACTGTCAACATTCAAGA 58


RESULT  36
AZ402172/c
LOCUS       AZ402172       96 bp    DNA             GSS       03-OCT-2000
DEFINITION  1M0169B24F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0169B24 F, DNA sequence.
ACCESSION   AZ402172
VERSION     AZ402172.1  GI:10517246
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 96)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0169  row: B  column: 24
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 96.
FEATURES             Location/Qualifiers
     source          1. .96
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0169B24"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       30 a      14 c      24 g      28 t
ORIGIN


  Query Match             55.5%;  Score 12.2;  DB 13;  Length 96;
  Best Local Similarity   82.4%;  Pred. No. 5.6e+04;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGA 17
```

```
          ||||||||||| |  ||
Db     71 TGACTGTGAACTTCAGA 55


RESULT  37
AI267734/c
LOCUS       AI267734       49 bp    mRNA            EST       17-NOV-1998
DEFINITION  ap62g07.x1 Stanley Frontal SN individual Homo sapiens cDNA clone
            IMAGE:2022204 similar to TR:Q28137 Q28137 MICROTUBULE-ASSOCIATED
            PROTEIN ;, mRNA sequence.
ACCESSION   AI267734
VERSION     AI267734.1  GI:3886901
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 49)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -40UP from Gibco
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .49
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2022204"
                     /clone_lib="Stanley Frontal SN individual"
                     /tissue_type="frontal lobe (see description)"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: brain; Vector: pCR2.1 (Invitrogen); Site_1:
                     EcoRI; Site_2: XhoI; Total RNA (purified with Trizol and
                     DNAsed before use) was reverse transcribed using a
                     modified oligo-dT primer containing RsaI and HindIII
                     sites.  Double- stranded cDNA was digested with RsaI,
                     resulting in blunt ended cDNA of an average 0.1-2 kb in
                     length.  Digested cDNA was split into two sets, one used
                     as is as the driver, the other set was split in half again
                     and each half linked to a different adaptor
                     (5'-TCGAGCGGCCGCCCGGGCAGGT-3' or 5'-
                     AGGGCGTGGTGCGGAGGGCGGT-3'), to be used as tester.
                     Subtraction was performed using the Clontech PCR Select
                     cDNA subtraction kit. 34 yo schizophrenic male (S-11)
                     subtracted by 28 yo mentally normal male (S-37). Tissues
                     were obtained from the Stanley Neuropathology Consortium
                     (www.stanleylab.org). Library constructed and subtracted
                     by Dr. Nancy Johnston [(410) 614-3918,
                     nlj@welchlink.welch.jhu.edu]."
BASE COUNT       23 a     12 c      8 g      6 t
ORIGIN


  Query Match           54.5%;  Score 12;  DB 10;  Length 49;
  Best Local Similarity  75.0%;  Pred. No. 5.9e+04;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGAT 20
          || ||||| | |||  ||||
Db     28 TGGCTGTGGATGTTTCAGAT 9


RESULT  38
AU107883/c
LOCUS       AU107883       50 bp    mRNA            EST       05-APR-2001
DEFINITION  AU107883 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            ZRV61130, mRNA sequence.
ACCESSION   AU107883
VERSION     AU107883.1  GI:13557405
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
            ,K., Suyama,A. and Sugano,S.
  TITLE     Fine Structural analysis of transcription start sites of human
            mRNAs using full-length enriched and 5'-end enriched cDNA libraries
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="ZRV61130"
                     /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT       8 a     13 c     12 g     16 t      1 others
ORIGIN


  Query Match           54.5%;  Score 12;  DB 10;  Length 50;
  Best Local Similarity  75.0%;  Pred. No. 5.9e+04;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGAT 20
          || ||| ||||  || ||||
Db     41 TGCCTGAGAACAGTCCAGAT 22


RESULT  39
TA346C05Q/c
LOCUS       TA346C05Q      50 bp    DNA             GSS       13-DEC-2000
DEFINITION  T. brucei sheared genomic DNA clone 346c05, reverse sequence,
            genomic survey sequence.
ACCESSION   AL496230
VERSION     AL496230.1  GI:11870136
KEYWORDS    GSS.
SOURCE      Trypanosoma brucei.
  ORGANISM  Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
            Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
            Melville,S.E., Rajandream,M.A. and Barrell,B.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
            project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
            Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
            nh1@sanger.ac.uk
COMMENT     Constructed at the Institute for Genomic Research (TIGR),
            Rockville, MD. Genomic DNA isolated from a cloned population of
            Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
            to give a tight size distribution (
```

4 kb). The v + i method used for the library construction is described in detail in Smith, H. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999).
Email: nelsayed@tigr.org
Details of T. brucei sequencing at the Sanger Centre are available at http://www.sanger.ac.uk/Projects/T_brucei/.

```
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Trypanosoma brucei"
                     /strain="TREU927"
                     /db_xref="taxon:5691"
                     /clone="346c05"
BASE COUNT       13 a      8 c     12 g     17 t
ORIGIN
```

```
  Query Match             54.5%;  Score 12;  DB 13;  Length 50;
  Best Local Similarity   75.0%;  Pred. No. 5.9e+04;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 TGACTGTGAACGTTCGAGAT 20
          ||||||||||  | ||  ||
Db     48 TGACTGTGAAAATACGCAAT 29
```

```
RESULT  40
AZ492525/c
LOCUS       AZ492525       56 bp    DNA              GSS       05-OCT-2000
DEFINITION  1M0326M09R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0326M09 R, DNA sequence.
ACCESSION   AZ492525
VERSION     AZ492525.1  GI:10665335
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 56)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0326  row: M  column: 09
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 56.
FEATURES             Location/Qualifiers
     source          1. .56
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0326M09"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       22 a      9 c      7 g     18 t
ORIGIN
```

```
  Query Match             54.5%;  Score 12;  DB 13;  Length 56;
  Best Local Similarity   75.0%;  Pred. No. 6.1e+04;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      3 ACTGTGAACGTTCGAGATGA 22
          ||| || | ||| ||||| |
Db     41 ACTTTGCATGTTAGAGATTA 22
```

```
RESULT  41
AA106075/c
LOCUS       AA106075       58 bp    mRNA             EST       04-FEB-1997
DEFINITION  ml87e04.r1 Stratagene mouse kidney (#937315) Mus musculus cDNA
            clone IMAGE:519006 5' similar to SW:ATP8_MOUSE P03930 ATP SYNTHASE
            PROTEIN 8 ;, mRNA sequence.
ACCESSION   AA106075
VERSION     AA106075.1  GI:1656124
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 58)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:312854
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28m13 rev1 ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .58
                     /organism="Mus musculus"
                     /strain="C57/Bl6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:519006"
                     /clone_lib="Stratagene mouse kidney (#937315)"
                     /sex="females"
                     /tissue_type="kidney"
                     /dev_stage="4 weeks"
                     /lab_host="SOLR (kanamycin resistant)"
```

```
                     /note="Organ: kidney; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                     Oligo dT. Average insert size: 1.0 kb; Uni-ZAP XR Vector;
                     ~5' adaptor sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor
                     sequence: 5' CTCGAGTTTTTTTTTTTTTTTTT 3'"
BASE COUNT       22 a      20 c       5 g      11 t
ORIGIN


  Query Match             54.5%;  Score 12;  DB 10;  Length 58;
  Best Local Similarity   75.0%;  Pred. No. 6.1e+04;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       3 ACTGTGAACGTTCGAGATGA 22
           | || ||| ||| | |||||
Db      27 AGTGGGAATGTTTGTGATGA 8


RESULT  42
BG514647
LOCUS       BG514647       64 bp    mRNA            EST       28-MAR-2001
DEFINITION  dad63b05.x1 Wellcome CRC pCS107 tropicalis egg Silurana tropicalis
            cDNA clone IMAGE:4463961 3', mRNA sequence.
ACCESSION   BG514647
VERSION     BG514647.1  GI:13485304
KEYWORDS    EST.
SOURCE      western clawed frog.
  ORGANISM  Silurana tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Silurana.
REFERENCE   1  (bases 1 to 64)
  AUTHORS   Clifton,S., Johnson,S.L., Blumberg,B., Song,J., Hillier,L., Pape,D.
            , Martin,J., Wylie,T., Underwood,K., Theising,B., Bowers,Y., Person
            ,B., Gibbons,M., Harvey,N., Ritter,E., Jackson,Y., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     WashU Xenopus EST project, 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Sandy Clifton, Ph.D.
            WashU Xenopus EST project, 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Library constructed by A. Zorn and J. Mason (Wellcome/CRC Institute
            ). DNA Sequencing by: Washington University Genome Sequencing
            Center
             Clone distribution: Xenopus clones from this library are available
            through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .64
                     /organism="Silurana tropicalis"
                     /db_xref="taxon:8364"
                     /clone="IMAGE:4463961"
                     /clone_lib="Wellcome CRC pCS107 tropicalis egg"
                     /tissue_type="egg"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Vector: pCS107; Site_1: NotI; Site_2: EcoRI; cDNAs
                     were oligo-dT primed and directionally cloned. Average
                     insert size 1.5 kb, range 0.5-4 kb. Library constructed by
                     A. Zorn and J. Mason (Wellcome/CRC Institute). "
BASE COUNT       20 a       9 c      14 g      21 t
ORIGIN


  Query Match             54.5%;  Score 12;  DB 11;  Length 64;
  Best Local Similarity   75.0%;  Pred. No. 6.3e+04;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       2 GACTGTGAACGTTCGAGATG 21
           | ||||| |||||  || ||
```

```
Db      30 GGCTGTGTACGTTACAGCTG 49


RESULT  43
TA111H12Q/c
LOCUS       TA111H12Q      66 bp    DNA             GSS       13-DEC-2000
DEFINITION  T. brucei sheared genomic DNA clone 111h12, reverse sequence,
            genomic survey sequence.
ACCESSION   AL460702
VERSION     AL460702.1  GI:11831978
KEYWORDS    GSS.
SOURCE      Trypanosoma brucei.
  ORGANISM  Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 66)
  AUTHORS   Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
            Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
            Melville,S.E., Rajandream,M.A. and Barrell,B.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
            project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
            Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
            nh1@sanger.ac.uk
COMMENT     Constructed at the Institute for Genomic Research (TIGR),
            Rockville, MD. Genomic DNA isolated from a cloned population of
            Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
            to give a tight size distribution (
            4 kb).  The v + i method used for the library construction is
            described in detail in Smith, H. and Venter, J.C. (Making small
            insert libraries for whole genome shotgun sequencing projects.  In
            Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
            Barrell, Oxford University Press, 1999).
            Email: nelsayed@tigr.org
            Details of T. brucei sequencing at the Sanger Centre are available
            at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES             Location/Qualifiers
     source          1. .66
                     /organism="Trypanosoma brucei"
                     /strain="TREU927"
                     /db_xref="taxon:5691"
                     /clone="111h12"
BASE COUNT       21 a      22 c       5 g      18 t
ORIGIN


  Query Match             54.5%;  Score 12;  DB 13;  Length 66;
  Best Local Similarity   75.0%;  Pred. No. 6.3e+04;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGAT 20
             ||| ||||  || |||| ||
Db      60 TGATTGTGGCCGATCGATAT 41


RESULT  44
CNS03NSG/c
LOCUS       CNS03NSG       68 bp    DNA             GSS       17-MAY-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence T7 end of clone
            040L07 of library G from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION   AL252457
VERSION     AL252457.1  GI:7973469
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis.
  ORGANISM  Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontidae; Tetraodon.
REFERENCE   1  (bases 1 to 68)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
            Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
```

```
            Weissenbach,J.
  TITLE     Charaterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 68)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
  TITLE     Human gene number estimate provided by genome wide analysis using
            Tetraodon nigroviridis DNA sequence
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 68)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000) to the EMBL/GenBank/DDBJ databases
COMMENT     This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES             Location/Qualifiers
     source          1. .68
                     /organism="Tetraodon nigroviridis"
                     /db_xref="taxon:99883"
                     /clone="040L07"
                     /clone_lib="G"
                     /note="Genoscope sequence ID : C0BG040CF04LP1~end : T7"
BASE COUNT       26 a     23 c     11 g      7 t      1 others
ORIGIN


  Query Match             54.5%;  Score 12;  DB 13;  Length 68;
  Best Local Similarity   75.0%;  Pred. No. 6.4e+04;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TGACTGTGAACGTTCGAGAT 20
           || ||| |  |||||||| |
Db      31 TGTCTGAGGGCGTTCGAGGT 12


RESULT  45
D86873/c
LOCUS       D86873         71 bp    DNA             GSS       07-FEB-1999
DEFINITION  Human exon sequence in 1.6Mb segment encompassing Down's syndrome
            region, exon, genomic survey sequence.
ACCESSION   D86873
VERSION     D86873.1  GI:1813387
KEYWORDS    GSS; exon trapping for Down's syndrome region.
SOURCE      Homo sapiens DNA, clone:E12-32.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 71)
  AUTHORS   Ohira,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-AUG-1996) to the DDBJ/EMBL/GenBank databases. Miki
            Ohira, Kazusa DNA Research Institute, Laboratory of Gene Structure
            1; 1532-3 Yanauchino, Kisarazu, Chiba 292, Japan
            (E-mail:oohira@kazusa.or.jp, Tel:+81-438-52-3932,
            Fax:+81-438-52-3931)
REFERENCE   2  (sites)
  AUTHORS   Ohira,M., Seki,N., Nagase,T., Ichikawa,H., Suzuki,E., Nomura,N. and
            Ohki,M.
  TITLE     Gene Identification in a 1.6-Mb region of the Down Syndrome Region
            on Chromosome 21
  JOURNAL   Unpublished (1996)
REFERENCE   3  (sites)
  AUTHORS   Ohira,M., Seki,N., Nagase,T., Suzuki,E., Nomura,N., Ohara,O.,
            Hattori,M., Sakaki,Y., Eki,T., Murakami,Y., Saito,T., Ichikawa,H.
            and Ohki,M.
  TITLE     Gene identification in the 1.6 -Mb of the down syndrome region on
            chromosome 21
  JOURNAL   Genome Res. (1997) In press
FEATURES             Location/Qualifiers
     source          1. .71
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /clone="E12-32"
                     /map="21q22.2"
     exon            1. .71
                     /note="trapped exon sequence"
BASE COUNT       14 a     26 c     12 g     19 t
ORIGIN


  Query Match             54.5%;  Score 12;  DB 13;  Length 71;
  Best Local Similarity   75.0%;  Pred. No. 6.5e+04;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       2 GACTGTGAACGTTCGAGATG 21
           ||| || |  |||| |||||
Db      42 GACGGTTAGAGTTCCAGATG 23


Search completed: November 29, 2001, 14:23:54
Job time: 8087 sec
```